

OM protein - protein search, using SW model

Run on: February 11, 2004, 14:35:52 ; Search time 8.64516 Seconds  
(without alignments)  
73.441 Million cell updates/sec

Title: US-10-050-611-1

Perfect score: 21

Sequence: 1 RCDA 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	4	13 AAR2315	Cell contact inhib
2	21	100.0	4	22 AAB6859	Transport molecule
3	21	100.0	4	23 AAE2833	Thrombo-spondin 1
4	21	100.0	4	23 AAE20157	Human thrombin pep
5	21	100.0	4	23 AAU78374	Thrombin peptide d
6	21	100.0	4	23 AAK50856	Thrombin receptor
7	21	100.0	5	13 AAR24517	Platelet anagons
8	21	100.0	5	20 AAU17781	Human thrombospond
9	21	100.0	5	22 AAB7600	Thrombin-induced p
10	21	100.0	6	11 AAR04871	Peptide from fibro
11	21	100.0	6	12 AAR11506	Cell attachment pr
12	21	100.0	7	23 ABR4835	Zinc finger protei
13	21	100.0	7	23 ABR4854	Zinc finger protei
14	21	100.0	7	23 ABR4857	Zinc finger protei
15	21	100.0	7	23 ABR4860	Zinc finger protei
16	21	100.0	7	23 ABR4863	Zinc finger protei
17	21	100.0	7	23 ABR4866	Zinc finger protei
18	21	100.0	7	23 ABR4869	Zinc finger protei
19	21	100.0	7	23 ABR4872	Zinc finger protei
20	21	100.0	7	23 ABR4883	Zinc finger protei
21	21	100.0	7	23 ABR4911	Zinc finger protei
22	21	100.0	7	23 ABR4914	Zinc finger protei
23	21	100.0	7	23 ABR4936	Zinc finger protei
24	21	100.0	7	23 ABR4943	Zinc finger protei
25	21	100.0	7	23 ABR4945	Zinc finger protei
26	21	100.0	7	23 ABR4946	Zinc finger protei
27	21	100.0	7	23 ABR4957	Zinc finger protei
28	21	100.0	7	23 ABR4963	Zinc finger protei
29	21	100.0	7	23 ABR4964	Zinc finger protei
30	21	100.0	7	23 ABR4965	Zinc finger protei
31	21	100.0	7	23 ABR4966	Zinc finger protei
32	21	100.0	7	23 ABR4967	Zinc finger protei
33	21	100.0	7	23 ABR4968	Zinc finger protei
34	21	100.0	7	23 ABR4969	Zinc finger protei
35	21	100.0	7	23 ABR4970	Zinc finger protei
36	21	100.0	7	23 ABR4971	Zinc finger protei
37	21	100.0	7	23 ABR4972	Zinc finger protei
38	21	100.0	7	23 ABR4973	Zinc finger protei
39	21	100.0	7	23 ABR4974	Zinc finger protei
40	21	100.0	7	23 ABR4975	Zinc finger protei
41	21	100.0	7	23 ABR4976	Zinc finger protei
42	21	100.0	7	23 ABR4977	Zinc finger protei
43	21	100.0	7	23 ABR4978	Zinc finger protei
44	21	100.0	8	19 ABR4979	Integrin receptor
45	21	100.0	8	24 ABR4980	Human FNfn10 FG 10

ALIGNMENTS

RESULT 1  
 AAR25315  
 ID : AAR25315 standard; peptide; 4 AA.  
 XX  
 AC AAR25315;  
 XX  
 DT 17-MAR-1993 (first entry)  
 XX  
 DE Cell contact inhibitor generic peptide #4.  
 XX  
 KW Cyclic peptide; cell contact inhibitor; hydrolytic enzyme.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 2  
 FT /label= Megly  
 XX  
 XX JF04264097-A.  
 XX  
 XX 18-SEP-1992.  
 XX  
 XX 16-FEB-1991; 91JP-0044386.  
 XX  
 XX 16-FEB-1991; 91JP-0044386.  
 XX  
 XX (ASAG ) ASAMI GLASS CO LTD.  
 XX  
 XX WPI; 1992-361922/44.  
 XX  
 PT Peptide derivs. as contact inhibitor for animal cells - comprise  
 PT synthesised cyclic peptide and have portion of aminoacid sequence  
 PT of arginine-N-methyl:glycine-aspartic acid  
 XX  
 XX Disclosure; Page 3; 6pp; Japanese.  
 XX  
 CC The sequences given in AAR25311-19 are cyclic peptides which act as  
 CC contact inhibitors of animal cells. They are resistant to  
 CC decomposition by hydrolytic enzymes and can be maintained at high  
 CC levels of activity for a long period in vivo. The peptides are  
 CC cyclic and may have 1-16 pref. 1-4 amino acids.  
 CC  
 XX Sequence 4 AA;  
 SQ  
 Query Match 100.0%; Score 21; DB 13; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDA 4  
 IIII  
 Db 1 RGDA 4  
 RESULT 2  
 AAB86859  
 ID : AAB86859 standard; peptide; 4 AA.  
 XX

AC AAB86859;  
 XX  
 DT 28-NOV-2001 (first entry)  
 XX  
 DE Transport molecule/ligand binding-associated peptide #5.  
 XX  
 KW Transport molecule; ligand; cancer treatment; autoimmune disease;  
 KW inflammation; infection.  
 XX  
 OS Synthetic.  
 XX  
 XX W0200168142-A1.  
 XX  
 XX 20-SEP-2001.  
 XX  
 XX 13-MAR-2001; 2001WO-EP02833.  
 XX  
 XX 13-MAR-2000; 2000DE-1012120.  
 XX  
 XX (KTB-T) KTB TUMORFORSCHUNGS GMBH.  
 XX  
 XX Kratz F;  
 XX  
 XX WPI; 2001-589998/66.  
 XX  
 XX New ligand, comprising therapeutic or diagnostic agent bonded  
 PT non-covalently with substance having high affinity to transport  
 PT molecule -  
 XX  
 XX Disclosure; Page 39; 74pp; German.  
 XX  
 CC This invention describes novel ligands which bind to transport molecules,  
 CC comprising a therapeutic and/or diagnostic agent (A) non-covalently  
 CC bonded via a linkage cleavable in vivo depending on pH and/or  
 CC enzymatically with a substance (B) having an association constant KA to a  
 CC transport molecule of above 10<sup>3</sup> M<sup>-1</sup>, is new. The medicaments are  
 CC especially useful for the treatment of cancers, autoimmune diseases,  
 CC acute and chronic inflammation and infections caused by viruses or  
 CC microorganisms. The diagnostic kits are useful for the detection of these  
 CC illnesses and for the detection of the transport molecule and/or its  
 CC distribution in vivo. The ligands have excellent solubility in the medium  
 CC at the site of action and are easy and inexpensive to covert into  
 CC adducts, as the interaction with the transport material is physical.  
 CC AAB86843-AAB86920 represent peptides used to illustrate the  
 CC method of the invention.  
 CC  
 XX Sequence 4 AA;  
 SQ  
 Query Match 100.0%; Score 21; DB 22; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDA 4  
 IIII  
 Db 1 RGDA 4

RESULT 3  
AAE28393  
ID AAE28393 standard; peptide; 4 AA.  
AC  
XX  
XX AAE28393;  
AC  
XX  
XX  
DT 27-DEC-2002 (first entry)  
DT  
XX  
XX Thrombo-spondin 1 RGD cell binding region.  
DE  
XX  
XX Tat region; nucleic acid-binding group; cell transfection system;  
KW gene therapy; cancer; thrombo-spondin 1.  
XX  
XX Unidentified.  
OS  
XX US6376248-B1.  
PN  
XX  
XX 23-APR-2002.  
PD  
XX  
XX 16-MAR-1998; 98US-0039780.  
PF  
XX  
XX 14-MAR-1997; 97US-0818200.  
PR  
XX  
XX (LIFE-) LIFE TECHNOLOGIES INC.  
PA  
XX  
XX Hawley-Nelson P, Ian J, Shih P, Jeesee JA, Schifferli KP;  
P1 Gebeyehu G, Ciccarone VC, Evans KL;  
P1 WPI; 2002-680647/73.  
DR  
XX  
XX New peptide comprising Tat sequence linked to nucleic acid-binding  
PT group, useful, e.g. in gene therapy, for improving cell-transfection  
PT efficiency -  
PT  
XX  
XX Example 1; Column 65; 108pp; English.  
PS  
XX  
XX The invention relates to a peptide comprising Tat sequence linked to  
CC nucleic acid-binding group. Peptides of the invention are used as  
CC components of a cell transfection system particularly for gene therapy  
CC (especially of cancer). The present sequence is thrombo-spondin 1 RGD  
CC cell binding region. This peptide is used in the exemplification of  
CC the invention.  
CC  
XX  
XX  
SQ Sequence 4 AA;  
Query Match 100.0%; Score 21; DB 23; Length 4;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX  
AC AAE20157;  
XX  
XX 18-JUN-2002 (first entry)  
XX  
XX  
XX Human thrombin peptide.  
DE  
XX  
XX Cartilage growth; cartilage repair; arthritic joint; traumatic injury;  
KW non-proteolytically activated thrombin receptor; NPAR; chondrocytes;  
KW therapy; implantation; thrombin peptide; human.  
XX  
XX Homo sapiens.  
OS  
XX W0200207748-A2.  
PN  
XX 31-JAN-2002.  
PD  
XX  
XX 19-JUL-2001; 2001WO-US22668.  
PF  
XX 20-JUL-2000; 2000US-219800P.  
PR  
XX  
XX (TEXA ) UNIV TEXAS SYSTEM.  
PA  
XX  
XX Carney DH, Crowther RS, Stienberg J, Bergmann J;  
P1 WPI; 2002-268953/31.  
DR  
XX  
XX Stimulating growth and repair of cartilage, useful for treating e.g.  
PT arthritis, by local administration of an agonist of non-proteolytically  
PT activated thrombin receptor -  
PT  
XX  
XX Claim 10; Page 25; 28pp; English.  
PS  
XX  
XX The invention relates to a method of stimulating growth and repair of  
CC cartilage. The method involves administering to the site, an agonist  
CC of non-proteolytically activated thrombin receptor (NPAR). The method  
CC is used in human or veterinary medicine for the treatment of arthritic  
CC joints and damage/loss of cartilage caused by traumatic injury. Also  
CC chondrocytes may be cultured in presence of NPAR agonist to provide  
CC cells for implantation at sites requiring growth/repair of cartilage.  
CC The present sequence is human thrombin peptide. The derivatives of  
CC thrombin peptide which serves as a NPAR agonist.  
CC  
XX  
XX  
SQ Sequence 4 AA;  
Query Match 100.0%; Score 21; DB 23; Length 4;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGD 4  
1111  
Db 1 RGD 4

QY 1 RGD 4  
1111  
Db 1 RGD 4

RESULT 4  
AAE20157  
ID AAE20157 standard; peptide; 4 AA.

RESULT 5  
AAU78374  
ID AAU78374 standard; peptide; 4 AA.

XX AAU78374;  
AC 18-JUN-2002 (first entry)  
DT Thrombin peptide derivative #1.  
DE Thrombin; osteopathic; receptor; agonist; bone growth stimulation;  
KW osteoinduction; farm animal; companion animal; laboratory animal;  
KW bone graft; segmental bone gap; bone void; non-union fracture.  
OS Synthetic.  
XX WO200205836-A2.  
XX 24-JAN-2002.  
XX 18-JUL-2001; 2001WO-US22641.  
XX 19-JUL-2000; 2000US-219300P.  
XX (TEXA ) UNIV TEXAS SYSTEM.  
XX Carney DH, Crowther RS, Simmons DJ, Yang J, Redin WR;  
XX WPI; 2002-303796/34.  
XX Stimulating bone growth at a site in a subject in need of  
PT osteoinduction, such as a site of bone graft, segmental bone gap, bone  
PT void or non-union structure, by administering agonist of activated  
PT thrombin receptor -  
XX  
XX Claim 9; Page 22; 27pp; English.  
XX  
CC The invention describes a method of stimulating bone growth at a site  
CC in a subject in need of osteoinduction. The method involves administering  
CC an agonist to stimulate bone growth at a site in a subject (e.g. a farm  
CC animal, companion animal or laboratory animal), in need of  
CC osteoinduction, such as the site in need of a bone graft in a subject, a  
CC segmental bone gap, a bone void or a non-union fracture. This sequence  
CC represents a thrombin peptide derivative obtained from a serine  
CC esterase that can stimulate or activate the non-protelytically  
CC activated thrombin receptor.  
XX  
SQ Sequence 4 AA;  
Query Match 100.0%; Score 21; DB 23; Length 4;  
Best Local Similarity 100.0%; Pred. No. 9.3e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ID AAM50856 standard; Peptide; 4 AA.  
XX  
XX AAM50856;  
AC 01-MAY-2002 (first entry)  
DT Thrombin receptor binding domain used for cardiac tissue repair.  
DE Thrombin receptor binding domain; thrombin; revascularisation;  
KW Thrombin receptor binding domain; thrombin; revascularisation;  
KW vascular occlusion; tissue repair; vulnery; vasotrophic; cardiac;  
KW angiogenesis; restenosis; therapy; human.  
XX  
XX Homo sapiens.  
XX WO200204008-A2.  
XX 17-JAN-2002.  
XX 12-JUL-2001; 2001WO-US21944.  
XX 12-JUL-2000; 2000US-217583P.  
XX (TEXA ) UNIV TEXAS SYSTEM.  
XX Carney DH;  
XX WPI; 2002-179665/23.  
XX Promoting cardiac tissue repair, stimulating revascularisation,  
PT stimulating vascular endothelial cell proliferation, and inhibiting  
PT vascular occlusion by using angiogenic thrombin derivative peptide -  
XX  
XX Claim 2; Page 19; 24pp; English.  
XX  
CC The present sequence is that of a thrombin receptor binding domain  
CC peptide that is used in a claimed method for promoting cardiac  
CC tissue repair. The method involves administering an angiogenic  
CC thrombin-derived peptide. The peptide comprises the present  
CC thrombin receptor binding domain together with a serine esterase  
CC conserved sequence (see AAM50857), or preferably a peptide (see  
CC AAM50858) which includes both these sequences. The thrombin-derived  
CC peptide is administered during or following cardiac surgery by  
CC injection into cardiac tissue, and may be formulated as a sustained  
CC release formulation. It is used in claimed methods of stimulating  
CC revascularisation, stimulating vascular endothelial cell  
CC proliferation, inhibiting vascular occlusion, and inhibiting  
CC restenosis following balloon angioplasty, in which case the  
CC peptide may be coated onto the catheter.  
XX  
SQ Sequence 4 AA;  
Query Match 100.0%; Score 21; DB 23; Length 4;  
Best Local Similarity 100.0%; Pred. No. 9.3e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 6  
AAM50856

QY 1 RGDA 4  
||||

Db 1 RGDA 4

RESULT 7  
AAR24517  
ID AAR24517 standard; Protein; 5 AA.  
XX  
AC AAR24517;  
XX  
DT 02-DEC-1992 (first entry)  
XX  
DE Platelet antagonist peptide 4.  
XX  
KW Clinical effect; antagonist.  
XX  
OS Synthetic.  
XX  
PN JP04134096-A.  
XX  
PD 07-MAY-1992.  
XX  
PE 21-SEP-1990; 90JP-0253849.  
XX  
PR 21-SEP-1990; 90JP-0253849.  
XX  
PA (SEGG) SEIKAGAKU KOGYO CO LTD.  
XX  
DR WPI; 1992-204525/25.  
XX  
PT New peptide(s) comprising arginine-glycine-asparagine and  
XX hyaluronic acid - useful as platelet antagonists with higher  
XX activity than arginine-glycine-asparagine-valine  
XX  
PS Disclosure; Page 5; 10pp; Japanese.  
XX  
CC The sequences given in AAR24514-8 are peptides which are useful as  
XX platelet antagonists. These peptides have higher activity than the  
XX conventional peptide of Arg-Gly-Asp-Val. These peptides have a  
XX clinical effect at a lower dose, dosage is 2.5-5.0 mg/kg/day.  
XX  
SQ Sequence 5 AA;  
Query Match 100.0%; Score 21; DB 13; Length 5;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4  
|||||  
Db 2 RGDA 5

RESULT 8  
AAV1781  
ID AAV1781 standard; peptide; 5 AA.  
XX  
AC AAV1781;  
XX

DT 12-AUG-1999 (first entry)  
XX  
DE Human thrombospondin-1 type III repeat peptide.  
XX  
KW Human; thrombospondin; HIV; infection; inhibition; chemokine;  
XX contractile.  
XX  
OS Homo sapiens.  
XX  
OS Synthetic.  
XX  
PN W09926649-A1.  
XX  
PD 03-JUN-1999.  
XX  
PF 24-NOV-1998; 98WO-US24905.  
XX  
PR 20-MAR-1998; 98US-0078873.  
XX  
PR 25-NOV-1997; 97US-0066294.  
XX  
PA (CORR) CORNELL RES FOUND INC.  
XX  
PI Crombie AR, Laurence JC, Nachman RL;  
XX  
DR WPI; 1999-370856/31.  
XX  
PT Suppressing infectivity of human immune deficiency virus  
XX  
PS Example 2; Page 33; 67pp; English.  
XX  
CC The present invention describes a method for suppressing infectivity of  
XX human immunodeficiency virus (HIV) by treating the virus, or its target  
XX cell, with a thrombospondin or thrombospondin analogue. Thrombospondin  
XX blocks binding of HIV to its cellular receptors. Thrombospondin or its  
XX analogues can be used to prevent infection by HIV, in both contraceptive  
XX and non-contraceptive compositions/devices. They are already known to  
XX reduce infectivity of some bacteria and protozoa. The present sequence  
XX represents a human thrombospondin-1 type III repeat peptide.  
XX  
SQ Sequence 5 AA;  
Query Match 100.0%; Score 21; DB 20; Length 5;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4  
|||||  
Db 2 RGDA 5

RESULT 9  
AAB72600  
ID AAB72600 standard; peptide; 5 AA.  
XX  
AC AAB72600;  
XX  
DT 09-MAY-2001 (first entry)  
XX

DE Thrombin-induced platelet activator antagonist #39.  
 XX  
 KW Platelet aggregation inhibitor; thrombin activation inhibitor;  
 KW Protease activated receptor 1; PAR1; platelet activation inhibitor;  
 KW thrombosis; acute coronary syndrome.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200112696-A1.  
 XX  
 PD 22-FEB-2001.  
 XX  
 PF 17-AUG-2000; 2000MO-US40669.  
 XX  
 PR 17-AUG-1999; 99US-0375808.  
 XX  
 PA (THRO-) THROMGEN INC.  
 XX  
 PI Schmaier AH, Hasan AAK;  
 DR WPI; 2001-226546/23.  
 DR  
 XX Inhibiting thrombin activation in human cell expressing protease  
 PT activated receptor 1 (PAR1), comprises contacting mixtures of thrombin  
 PT and human cell expressing PAR1, with a peptide that inhibits platelet  
 PT activation -  
 XX  
 PS Claim 8; Page 26; 49pp; English.  
 XX  
 CC The present invention relates to a method for inhibiting thrombin  
 CC activation in a human cell expressing protease activated receptor 1  
 CC (PAR1). The method involves using peptides (e.g. the present peptide)  
 CC that inhibit platelet activation. The method is useful for preventing  
 CC thrombosis and platelet aggregation. The method can be used for patients  
 CC with acute coronary syndromes (e.g. crescendo angina, myocardial  
 CC infarction) and for individuals who have acute coronary syndromes and  
 CC receive percutaneous transluminal coronary angioplasty with an artificial  
 CC stent placement.  
 CC  
 XX Sequence 5 AA;  
 SQ  
 Query Match 100.0%; Score 21; DB 22; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGD 4  
 DB 1 RGD 4  
 QY 1 RGD 4  
 DB 1 RGD 4  
 RESULT 10  
 AAR04871  
 ID AAR04871 standard; peptide; 6 AA.  
 XX  
 AC AAR04871;  
 XX  
 DE 25-MAR-2003 (updated)

DT 25-SEP-1989 (first entry)  
 XX  
 DE Peptide from fibrinectin.  
 XX  
 KW Fibrinectin; cell attachment; cell detachment; fermentation; therapy.  
 XX  
 OS synthetic.  
 XX  
 PN US4879237-A.  
 XX  
 PD 07-NOV-1989.  
 XX  
 PF 24-MAY-1985; 85US-0736078.  
 XX  
 PR 24-MAY-1985; 85US-0736078.  
 XX  
 PA (LJOL-) LA JOLLA CANCER RES FOUND.  
 XX  
 PI Ruoslahti EI, Hayman EG, Pierschbacher MD;  
 DR WPI; 1990-154405/20.  
 DR  
 XX Synthetic peptide(s) from fibrinectin- used in control of cell attachment  
 PT and detachment  
 PT  
 PS Claim 1; page 10; 13pp; English.  
 XX  
 CC This polypeptide mediates the attachment of animal cells to substrates.  
 CC The substrate (1) is contacted with cells and with a soln. contg. this  
 CC polypeptide. This attachment can be prevented in addition to detaching  
 CC the cells from (1) once attached. Applications are in eg fermentation,  
 CC cell line prepn., diagnosis and therapy.  
 CC (Updated on 25-MAR-2003 to correct PR field.)  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 CC  
 XX Sequence 6 AA;  
 SQ  
 Query Match 100.0%; Score 21; DB 11; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGD 4  
 DB 2 RGD 5  
 QY 1 RGD 4  
 DB 2 RGD 5  
 RESULT 11  
 AAR11506  
 ID AAR11506 standard; Protein; 6 AA.  
 XX  
 AC AAR11506;  
 XX  
 DT 12-JUN-1991 (first entry)  
 XX  
 DE Cell attachment promoting peptide.  
 XX  
 KW Fibrin; aggregation.



Query Match 100.0%; Score 21; DB 23; Length 7;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4  
|||||  
Db 1 RGDA 4

RESULT 13

ABP48594  
ID ABP48594 standard; Peptide; 7 AA.  
AC ABP48594;  
XX  
DT 28-AUG-2002 (first entry)  
XX  
DE Zinc finger protein related peptide motif SEQ ID NO:670.  
XX  
KM Zinc finger protein; ZFP; DNA binding protein; zinc finger.  
XX  
OS Homo sapiens.  
XX  
OS Synthetic.  
XX  
EN W0200242459-A2.  
XX  
PD 30-MAY-2002.  
XX  
PF 20-NOV-2001; 2001WO-US43438.  
XX  
PR 20-NOV-2000; 2000US-0716637.  
XX  
PA (SANG-) SANGAMO BIOSCIENCES INC.  
XX  
PI Liu Q;  
XX  
PI WPI; 2002-500284/53.  
XX  
PT New zinc finger protein that binds to target site, useful in studying  
PT gene function and for human therapeutics and plant engineering,  
PT comprises first, second and third zinc fingers, ordered from N- to  
PT C-terminus -  
XX  
PS Example 1; Page 40; 81pp; English.

CC The present invention describes a zinc finger protein (I) that binds to  
CC a target site, comprising a first (F1), a second (F2), and a third (F3)  
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the  
CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),  
CC and a third (S3) target sub-site. Also described are: (1) a polypeptide  
CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and  
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that  
CC it binds to the S1 target sub-site, selecting the F2 zinc finger such  
CC that it binds to the S2 target sub-site, and selecting the F3 zinc  
CC finger such that it binds to the S3 target sub-site, thus designing (I)  
CC that binds to a target site. (I) is useful for recognition of triplet  
CC target sub-sites having the nucleotide G in the 5'-most position of the

CC sub-site. (I) is useful in studying gene function, and for human  
CC therapeutics and plant engineering. (I), (II) or (III) is useful in  
CC therapeutic methods to modulate the expression of a target region within  
CC a subject, in diagnostic methods for sequence specific detection of  
CC target nucleic acid in a sample, and in assays to determine the  
CC phenotype and function of gene expression. (I) has improved affinity  
CC and specificity for their target sequences, as well as enhanced  
CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP1230  
CC represent DNA target sequences and zinc finger peptides which are given  
CC in the exemplification of the present invention.  
XX

SEQ Sequence 7 AA.

QY 1 RGDA 4  
|||||  
Db 1 RGDA 4

RESULT 14

ABP48597  
ID ABP48597 standard; Peptide; 7 AA.  
XX  
AC ABP48597;  
XX  
DT 28-AUG-2002 (first entry)  
XX  
DE Zinc finger protein related peptide motif SEQ ID NO:671.  
XX  
KM Zinc finger protein; ZFP; DNA binding protein; zinc finger.  
XX  
OS Homo sapiens.  
XX  
OS Synthetic.  
XX  
EN W0200242459-A2.  
XX  
PD 30-MAY-2002.  
XX  
PF 20-NOV-2001; 2001WO-US43438.  
XX  
PR 20-NOV-2000; 2000US-0716637.  
XX  
PA (SANG-) SANGAMO BIOSCIENCES INC.  
XX  
PI Liu Q;  
XX  
PI WPI; 2002-500284/53.  
XX  
PT New zinc finger protein that binds to target site, useful in studying  
PT gene function and for human therapeutics and plant engineering,  
PT comprises first, second and third zinc fingers, ordered from N- to  
PT C-terminus -  
XX  
PS Example 1; Page 40; 81pp; English.

XX The present invention describes a zinc finger protein (I) that binds to  
 CC a target site, comprising a first (F1), a second (F2), and a third (F3)  
 CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the  
 CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),  
 CC and a third (S3) target sub-site. Also described are: (I) a polypeptide  
 CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and  
 CC (3) designing (M) (I) involves selecting the F1 zinc finger such that  
 CC it binds to the S1 target sub-site, selecting the F2 zinc finger such  
 CC that it binds to the S2 target sub-site, and selecting the F3 zinc  
 CC finger such that it binds to the S3 target sub-site, thus designing (I)  
 CC that binds to a target site. (I) is useful for recognition of triplet  
 CC target sub-sites having the nucleotide G in the 5'-most position of the  
 CC sub-site. (I) is useful in studying gene function, and for human  
 CC therapeutics and plant engineering. (I), (II) or (III) is useful in  
 CC therapeutic methods to modulate the expression of a target region within  
 CC a subject, in diagnostic methods for sequence specific detection of  
 CC target nucleic acid in a sample, and in assays to determine the  
 CC phenotype and function of gene expression. (I) has improved affinity  
 CC and specificity for their target sequences, as well as enhanced  
 CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230  
 CC represent DNA target sequences and zinc finger peptides which are given  
 CC in the exemplification of the present invention.

XX Sequence 7 AA;  
 SQ

Query Match 100.0%; Score 21; DB 23; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGD4 4  
 ||||  
 DB 1 RGD4 4

RESULT 15  
 ABP48600  
 ID ABP48600 standard; Peptide; 7 AA.  
 XX  
 AC ABP48600;  
 XX  
 DT 28-AUG-2002 (first entry)  
 XX  
 DE Zinc finger protein related peptide motif SBQ ID NO:672.  
 XX  
 XX Zinc finger protein; ZFP; DNA binding protein; zinc finger.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PW W0200242459-A2.  
 XX  
 PD 30-MAY-2002.  
 XX  
 PF 20-NOV-2001; 2001WO-US43438.  
 XX  
 PR 20-NOV-2000; 2000US-0716637.

XX (SANG-) SANGMO BIOSCIENCES INC.  
 PA  
 XX  
 PI Liu Q;  
 XX  
 DR WPI; 2002-500284/53.  
 XX  
 PT New zinc finger protein that binds to target site, useful in studying  
 PT gene function and for human therapeutics and plant engineering,  
 PT comprises first, second and third zinc fingers, ordered from N- to  
 PT C-terminus -  
 PS  
 XX Example 1; Page 40; 81pp; English.  
 XX  
 CC The present invention describes a zinc finger protein (I) that binds to  
 CC a target site, comprising a first (F1), a second (F2), and a third (F3)  
 CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the  
 CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),  
 CC and a third (S3) target sub-site. Also described are: (I) a polypeptide  
 CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and  
 CC (3) designing (M) (I) involves selecting the F1 zinc finger such that  
 CC it binds to the S1 target sub-site, selecting the F2 zinc finger such  
 CC that it binds to the S2 target sub-site, and selecting the F3 zinc  
 CC finger such that it binds to the S3 target sub-site, thus designing (I)  
 CC that binds to a target site. (I) is useful for recognition of triplet  
 CC target sub-sites having the nucleotide G in the 5'-most position of the  
 CC sub-site. (I) is useful in studying gene function, and for human  
 CC therapeutics and plant engineering. (I), (II) or (III) is useful in  
 CC therapeutic methods to modulate the expression of a target region within  
 CC a subject, in diagnostic methods for sequence specific detection of  
 CC target nucleic acid in a sample, and in assays to determine the  
 CC phenotype and function of gene expression. (I) has improved affinity  
 CC and specificity for their target sequences, as well as enhanced  
 CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230  
 CC represent DNA target sequences and zinc finger peptides which are given  
 CC in the exemplification of the present invention.

XX Sequence 7 AA;  
 SQ

Query Match 100.0%; Score 21; DB 23; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGD4 4  
 ||||  
 DB 1 RGD4 4

Search completed: February 11, 2004, 14:53:24  
 Job time : 10.6452 secs

OK protein - protein search, using sw model

Run on: February 11, 2004, 14:49:07; Search time 2.70968 Seconds  
(without alignments)  
141.963 Million cell updates/sec

Title: US-10-050-611-1

Perfect score: 21

Sequence: 1 RCDA 4

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

1: p1r1: \*  
2: p1r2: \*  
3: p1r3: \*  
4: p1r4: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	19	2 A34467	36K microfilbril-as
2	21	100.0	39	2 A34453	decorin - leech (
3	21	100.0	45	2 G82832	hypothetical prote
4	21	100.0	49	2 S70393	hypothetical prote
5	21	100.0	52	2 S19623	ornatin C - leech
6	21	100.0	57	2 E70535	hypothetical prote
7	21	100.0	68	2 AG3217	hypothetical prote
8	21	100.0	74	2 S62570	60S ribosomal prot
9	21	100.0	76	2 I39505	tcp RNA-binding pr
10	21	100.0	79	2 B90870	hypothetical prote
11	21	100.0	79	2 G85748	unknown protein en
12	21	100.0	79	2 E64864	ydaQ protein - Esc

13	21	100.0	80	2 S68677	cytochrome c551 -
14	21	100.0	88	2 H82662	conserved hypothet
15	21	100.0	89	2 I66553	cell surface glyco
16	21	100.0	90	2 E82562	hypothetical prote
17	21	100.0	93	2 AH0620	probable prephage
18	21	100.0	93	2 E82696	hypothetical prote
19	21	100.0	96	2 G84240	hypothetical prote
20	21	100.0	96	2 D83771	hypothetical prote
21	21	100.0	97	2 A71054	ribosomal protein
22	21	100.0	97	2 C75089	ribosomal protein
23	21	100.0	97	2 E82962	hypothetical prote
24	21	100.0	98	2 S01566	hypothetical prote
25	21	100.0	100	2 T30673	hypothetical prote
26	21	100.0	102	2 E75273	conserved hypothet
27	21	100.0	103	2 F70976	hypothetical prote
28	21	100.0	104	2 B72538	probable acylphosp
29	21	100.0	107	2 F90230	partial transposas
30	21	100.0	108	2 T51207	hypothetical prote
31	21	100.0	110	2 AC2787	conserved hypothet
32	21	100.0	110	2 E97566	hypothetical prote
33	21	100.0	115	2 S14024	hypothetical prote
34	21	100.0	115	2 C82479	hypothetical prote
35	21	100.0	116	2 D71832	ribosomal protein
36	21	100.0	116	2 D64681	ribosomal protein
37	21	100.0	117	2 B81255	50S ribosomal prot
38	21	100.0	121	2 I35719	pinQ protein - Esc
39	21	100.0	123	2 H75059	hypothetical prote
40	21	100.0	124	2 D84319	30S ribosomal prot
41	21	100.0	124	2 S62816	ribosomal protein
42	21	100.0	124	2 T03574	hypothetical prote
43	21	100.0	126	2 C86883	50S ribosomal prot
44	21	100.0	126	2 B72621	hypothetical prote
45	21	100.0	126	2 I37063	hypothetical prote

#### ALIGNMENTS

RESULT 1  
A34467  
36K microfilbril-associated protein - pig (fragment)  
C/Species: Sus scrofa domestica (domestic pig)  
C/Date: 08-Jun-1990 #sequence\_revision 08-Jun-1990 #text\_change 18-Jun-1993  
C/Accession: A34467  
R/Kobayashi, R.; Tashima, Y.; Masuda, H.; Shozawa, T.; Numata, Y.; Miyasoh, K.;  
Hayakawa, T.  
J. Biol. Chem. 264, 17437-17444, 1989  
A/Title: Isolation and characterization of a new 36-kDa microfilbril-associated  
glycoprotein from porcine aorta.  
A/Reference number: A34467; NUID:9008913; PMID:2793866  
A/Accession: A34467  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 1-19 <NO>  
Query Match 100.0%; Score 21; DB 2; Length 19;  
Best Local Similarity 100.0%; Pred. No. 60;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4  
|||||  
Db 5 RGDA 8

## RESULT 2

A36453  
decorin - leech (*Macrobdella decora*)  
C/Species: *Macrobdella decora*  
C/Date: 08-Mar-1991 #sequence\_revision 08-Mar-1991 #text\_change 30-Sep-1993  
C/Accession: A36453  
R/Seymour, J.L.; Henzel, W.J.; Nevins, B.; Stults, J.T.; Lazarus, R.A.  
J. Biol. Chem. 265, 10143-10147, 1990  
A>Title: Decorin. A potent glycoprotein IIB-IIIA antagonist and platelet aggregation inhibitor from the leech *Macrobdella decora*.  
A/Reference number: A36453; PMID:90277628; PMID:2351655  
A/Accession: A36453  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 1-39 <SEV>

Query Match 100.0%; Score 21; DB 2; Length 39;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4  
|||||  
Db 31 RGDA 34

## RESULT 3

G82812  
hypothetical protein XF0386 [imported] - *Xylella fastidiosa* (strain 9a5c)  
C/Species: *Xylella fastidiosa*  
C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C/Accession: G82812  
R/Anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequencing and Analysis, Sao Paulo, Brazil.  
Nature 406, 151-157, 2000  
A>Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.  
A/Reference number: A82515; PMID:20365717; PMID:10910347  
A/Note: for a complete list of authors see reference number A59328 below  
A/Accession: G82812  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-45 <STM>  
A/Cross-references: GB:AE003890; GB:AE003849; NID:9105215; PIDN:AA83196.1;  
GSPDB:Q00128; XFSC:XF0386  
A/Experimental source: strain 9a5c  
R/Simpson, A.J.G.; Reinach, F.C.; Aruda, P.; Abreu, F.A.; Acencio, M.; Alvarado, R.; Alves, L.M.C.; Araya, J.E.; Bala, G.S.; Baptista, C.S.; Barros, M.H.; Bonaccorsi, E.D.; Bordin, S.; Bove, J.M.; Brites, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carreir, H.; Colauto, N.B.; Colombo, C.; Costa, F.F.; Costa, M.C.R.; Costa-Veto, C.M.; Coutinho, L.L.;

Cristofani, M.; Dias-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000

A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fromme, M.; Furian, L.R.; Garnier, M.; Goldman, G.H.; Goldman, M.H.S.; Gomes, S.L.; Gruber, A.; Ho, P.L.; Hohnselt, J.D.; Junqueira, M.L.; Kemper, E.L.; Kitejima, U.P.; Krieger, J.E.; Kuremaa, E.B.; Laigret, F.; Lampais, M.R.; Leite, L.C.C.; Lemos, E.G.M.; Lemos, M.V.F.; Lopes, S.A.; Lopes, C.R.; Machado, J.A.; Machado, M.A.; Madeira, A.M.B.N.; Madella, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.A.L.

A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; Monteiro-Vitorello, C.B.; Moon, D.H.; Nagai, M.A.; Nascimento, A.L.T.O.; Netto, L.E.S.; Nham Jr., A.; Nobrega, E.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.; Paris, A.; Pelxoto, B.R.; Pereira, G.A.G.; Pereira Jr., H.A.; Pasquero, J.B.; Queaglio, R.B.; Roberto, P.G.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki, H.E.

A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silva, J.F.; Silvestri, M.L.Z.; Siqueira, W.J.; de Souza, A.A.; de Souza, A.P.; Terenzi, M.F.; Trufi, D.; Tsai, S.M.; Tsunako, M.H.; Vallada, H.; Van Slyke, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Zago, M.A.; Zatz, M.; Weidants, J.; Setubal, J.C.

A/Reference number: A59328  
A/Contents: annotation  
C/Genetics:  
A/Accession: XF0386

Query Match 100.0%; Score 21; DB 2; Length 45;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4  
|||||  
Db 19 RGDA 22

## RESULT 4

S70093  
hypothetical protein (orf49) - *Amycolatopsis methanolica*  
C/Species: *Amycolatopsis methanolica*  
C/Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 07-May-1999  
C/Accession: S70093  
R/Vrijbloed, J.W.; Jelinkova, M.; Hessels, G.I.; Dijkhuizen, L.  
Mol. Microbiol. 18, 21-31, 1995  
A>Title: Identification of the minimal replicon of plasmid pME300 of the methylotrophic actinomycete *Amycolatopsis methanolica*.  
A/Reference number: S70087; PMID:96154938; PMID:8596458  
A/Accession: S70093  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-49 <VR1>  
A/Cross-references: EMBL:J36679  
C/Genetics:  
A/Start codon: GTG

Query Match 100.0%; Score 21; DB 2; Length 49;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 4/ Conservative 0/ Mismatches 0/ Indels 0/ Gaps 0/

QY 1 RGDA 4  
|||||

DB 23 RGDA 26

RESULT 5  
S19623  
ornatin C - leech (Placobdella ornata)  
C/Species: Placobdella ornata  
C/Date: 19-Mar-1997 #sequence\_revision 13-Mar-1997 #text\_change 24-Jul-1997  
C/Accession: S19623  
R/Mazur, P.; Henzel, W.J.; Seymour, J.L.; Lazarus, R.A.  
Eur. J. Biochem. 202, 1073-1082, 1991  
A/Title: Ornatin: potent glycoprotein IIb-IIIa antagonists and platelet aggregation inhibitors from the leech Placobdella ornata.  
A/Reference number: S19566; PMID:92111479; PMID:1765068  
A/Accession: S19623  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 1-92 <MAZ>

Query Match 100.0%; Score 21; DB 2; Length 52;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 4/ Conservative 0/ Mismatches 0/ Indels 0/ Gaps 0/

QY 1 RGDA 4  
|||||

DB 42 RGDA 45

RESULT 6  
E70535  
hypothetical protein RV0666 - Mycobacterium tuberculosis (strain H37RV)  
C/Species: Mycobacterium tuberculosis  
C/Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C/Accession: E70535  
R/Ole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.V.; Eigmeier, K.; Gasp, S.; Barry III, C.E.; Tekle, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Felkell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.; Rajadaram, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.  
A/Reference number: A70500; PMID:98295987; PMID:9634230  
A/Accession: E70535  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-97 <COL>  
A/Cross-references: GB:Z95972; GB:AJ123456; NID:93261790; PIDN:CAM09391.1; PID:E319190; PID:92143295  
A/Experimental source: strain H37RV

C/Genetics:  
A/Gene: RV0666

Query Match 100.0%; Score 21; DB 2; Length 57;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 4/ Conservative 0/ Mismatches 0/ Indels 0/ Gaps 0/

QY 1 RGDA 4  
|||||

DB 24 RGDA 27

RESULT 7  
AG3217  
hypothetical protein Atu5470 [imported] - Agrobacterium tumefaciens (strain C58, Dupont) plasmid AT  
C/Species: Agrobacterium tumefaciens  
C/Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
C/Accession: AG3217  
R/Wood, D.W.; Sebubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.; Kitejima, J.P.; Okura, V.K.; Almeida Jr., N.F.; Zhou, Y.; Bovee Sr., D.; Chapman, P.; Clendinning, J.; Deatherage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kucyavin, T.; Levy, R.; Li, M.; McCelland, E.; Palmeri, A.; Raymond, C.; Rouse, G.; Saepthimachak, C.; Wu, Z.; Gordon, D.; Eisen, J.A.; Paulsen, I.; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A/Authors: too, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, B.; Liao, L.; Kim, S.; Hendrick, C.; Zhao, Z.; Dolan, M.; Tinney, S.V.; Tomb, J.; Gordon, M.P.; Olson, M.V.; Nester, E.W.  
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A/Reference number: AB2577; PMID:21608550; PMID:11743193  
A/Accession: AG3217  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-68 <KUR>  
A/Cross-references: GB:AE006687; PIDN:AA146157.1; PID:g17743927; GSPDB:GN00188  
A/Experimental source: strain C58 (Dupont)  
C/Genetics:  
A/Gene: Atu5470  
A/Genome: plasmid

Query Match 100.0%; Score 21; DB 2; Length 68;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 4/ Conservative 0/ Mismatches 0/ Indels 0/ Gaps 0/

QY 1 RGDA 4  
|||||

DB 36 RGDA 39

RESULT 8  
S62570  
60S ribosomal protein l38 - fission yeast (Schizosaccharomyces pombe)  
N/Alternate names: protein SPAC30D11.1  
C/Species: Schizosaccharomyces pombe  
C/Date: 06-Dec-1996 #sequence\_revision 06-Dec-1996 #text\_change 11-Jan-2000

C/Accession: S62570; T38587  
 R/Pearson, D.; Churcher, C.M.  
 Submitted to the EMBL Data Library, November 1995  
 A/Reference number: S62559  
 A/Accession: S62570  
 A/Molecule type: DNA  
 A/Residues: 1-74 <PEP>  
 A/Cross-references: EMBL:Z67961; NID:q1065887; PIDD:CAA91898.1; PID:q1065899  
 R/Pearson, D.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
 Submitted to the EMBL Data Library, November 1995  
 A/Reference number: Z21801  
 A/Accession: T38587  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-74 <PEP>  
 A/Cross-references: EMBL:Z67961; PIDD:CAA91898.1; GSPDB:GN00066;  
 SPDB:SPAC30D11.12  
 A/Experimental source: strain 97zh-; cosmid c30D11  
 C/Genetics:  
 A/Genes: FP138-2; SPAC30D11.12  
 A/Map position: 1L  
 A/Intons: 1/3; 64/1  
 C/Superfamily: rat ribosomal protein L38  
 C/Keywords: cytosol; protein biosynthesis; ribosome  
  
 Query Match 100.0%; Score 21; DB 2; Length 74;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 1 RGDA 4  
 ||||  
 DB 17 RGDA 20  
  
 RESULT 9  
 139905  
 C/Species: Bacillus pumilus  
 C/Date: 19-Jul-1996 #sequence\_revision 19-Jul-1996 #text\_change 15-Oct-1999  
 C/Accession: 139905  
 R/Hofman, R.J.; Gollnick, P.  
 J. Bacteriol. 177, 839-842, 1995  
 A/Title: The mtrB gene of Bacillus pumilus encodes a protein with sequence and functional homology to the trp RNA-binding attenuation protein (TRAP) of Bacillus subtilis.  
 A/Reference number: 139904; NID:95138053; PMID:7836324  
 A/Accession: 139905  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-76 <RES>  
 A/Cross-references: GB:L37879; NID:9598076; PIDD:AAA67544.1; PID:9598078  
 C/Genetics:  
 A/Genes: mtrB  
  
 Query Match 100.0%; Score 21; DB 2; Length 76;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4  
 ||||  
 DB 58 RGDA 61  
  
 RESULT 10  
 B90870  
 C/Species: Escherichia coli  
 C/Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
 C/Accession: B90870  
 R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Ohtsubo, E.; Nakayama, K.; Murata, T.; Tanaka, M.; Tobe, I.; Iida, T.; Takami, H.; Honda, T.; Sasakawa, C.; Ogatawara, N.; Yasunaga, T.; Kohara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
 DNA Res. 8, 11-22, 2001  
 A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12.  
 A/Reference number: A99629; NID:21156231; PMID:11258796  
 A/Accession: B90870  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-79 <HAV>  
 A/Cross-references: GB:BA000007; PIDD:BA83535.1; PID:q13361395; GSPDB:GN00154  
 A/Experimental source: strain O157:H7, substrain RIMD 0509552  
 C/Genetics:  
 A/Genes: ECa1930  
  
 Query Match 100.0%; Score 21; DB 2; Length 79;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 1 RGDA 4  
 ||||  
 DB 5 RGDA 8  
  
 RESULT 11  
 G85748  
 C/Species: Escherichia coli  
 C/Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
 C/Accession: G85748  
 R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, G.F.; Evans, P.S.; Gregor, J.; Kirkpatrick, H.A.; Posfai, G.; Hackett, J.; Kink, S.; Boutin, A.; Shao, Y.; Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dialante, E.; Potamousis, K.; Apodaca, J.; Anantharaman, T.S.; Lin, J.; Yen, G.; Schwartz, D.C.; Welch, R.A.; Blattner, F.R.  
 Nature 409, 529-533, 2001  
 A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A/Reference number: A85480; NID:21074935; PMID:11206551  
 A/Accession: G85748  
 A/Status: preliminary  
 A/Molecule type: DNA

A/Residues: 1-79 <STO>  
A/Cross-references: GB:AE005174; NID:q12515406; PIDN:AA056451.1; GSPDB:GN00145;  
UMCP:22414  
A/Experimental source: strain O157:H7, substrain EDL533  
C/Genetics:  
A/Gene: ydaQ

Query Match 100.0%; Score 21; DB 2; Length 79;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4  
||||  
DB 5 RGDA 8

RESULT 12  
ydaQ protein - Escherichia coli (strain K-12)  
C/Species: Escherichia coli  
C/Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002  
C/Accession: E64884  
R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.;  
Riley, M.; Collado-Vides, J.; Glasner, J.D.; Rode, C.K.; Mayhew, G.F.; Gregor,  
J.; Davis, N.W.; Kirkpatrick, H.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A/Title: The complete genome sequence of Escherichia coli K-12.  
A/Reference number: A64720; MUID:97426617; PMID:9278503  
A/Accession: E64884  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-79 <BLAT>  
A/Cross-references: GB:AE000232; GB:U00096; NID:q1787600; PIDN:AACT4428.1;  
PID:q1787608; UMCP:51346  
A/Experimental source: strain K-12, substrain MG1655  
C/Genetics:  
A/Gene: ydaQ

Query Match 100.0%; Score 21; DB 2; Length 79;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4  
||||  
DB 5 RGDA 8

RESULT 13  
S68677  
cytochrome c551 - Chromatium vinosum  
C/Species: Chromatium vinosum  
C/Date: 25-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 04-Mar-2000  
C/Accession: S68677  
R/Samyn, B.; de Smet, L.; van Driessche, G.; Meyer, T.E.; Bartsch, R.G.;  
Cusenovitch, M.A.; van Beeumen, J.J.  
Eur. J. Biochem. 236, 689-696, 1996

A/Title: A high-potential soluble cytochrome c-551 from the purple phototrophic  
bacterium Chromatium vinosum is homologous to cytochrome c(8) from denitrifying  
pseudomonas.  
A/Reference number: S68677; MUID:96193682; PMID:8612646  
A/Accession: S68677  
A/Molecule type: protein  
A/Residues: 1-80 <SAV>  
A/Experimental source: strain D  
C/Superfamily: cytochrome c6; cytochrome c6 homology  
C/Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein;  
oxidative phosphorylation  
F1:77/Domain: cytochrome c6 homology <CYC>  
F10:13/Binding site: heme (Cys) (covalent) #status predicted  
F14:59/Binding site: heme iron (His, Met) (axial ligands) #status predicted

Query Match 100.0%; Score 21; DB 2; Length 80;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4  
||||  
DB 33 RGDA 36

RESULT 14  
H82662  
conserved hypothetical protein XF1562 [Imported] - Xylella fastidiosa (strain  
945C)  
C/Species: Xylella fastidiosa  
C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C/Accession: H82662  
R/Anonymous, The Xylella fastidiosa Consortium of the Organization for  
Nucleotide Sequencing and Analysis, Sao Paulo, Brazil.  
Nature 406, 151-157, 2000  
A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A/Reference number: A82515; MUID:20365717; PMID:10910347  
A/Note: for a complete list of authors see reference number A59328 below  
A/Accession: H82662  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-86 <STN>  
A/Cross-references: GB:AE003986; GB:AE003849; NID:q910606; PIDN:AAFE94371.1;  
GSPDB:GN00128; XFSC:XF1562  
A/Experimental source: strain 945C  
R/Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.;  
Alvarenga, R.; Alves, L.M.C.; Araya, J.E.; Bala, G.S.; Baptista, C.S.; Barros,  
M.H.; Bonaccorsi, E.D.; Bordin, S.; Bove, J.M.; Briones, M.R.S.; Bueno, M.R.P.;  
Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carter, H.; Colauto, N.B.;  
Colombo, C.; Costa, F.F.; Costa, M.C.R.; Costa-Neto, C.M.; Coutinho, L.L.;  
Cristofani, M.; Dias-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.;  
Ferreira, A.V.S.  
submitted to GenBank, June 2000  
A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco,  
M.C.; Frohme, M.; Furlan, L.R.; Garnier, M.; Goldman, G.H.; Goldman, M.H.S.;  
Gomes, S.L.; Gruber, A.; Ho, P.L.; Honzels, J.D.; Iungueira, M.L.; Kemper,  
E.; Kitajima, J.P.; Krieger, J.B.; Kurmae, E.; Laigret, F.; Lambais, M.R.;  
Lette, L.C.C.; Lemos, E.G.M.; Lemos, M.V.F.; Lopes, S.A.; Lopes, C.R.; Machado,

J.A.; Machado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.A.L.  
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; Monteiro-Vitorello, C.B.; Moon, D.H.; Nagai, M.A.; Nascimento, A.L.T.O.; Netto, L.E.S.; Nham Jr., A.; Nobrega, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.; Paris, A.; Peixoto, B.R.; Pereira, G.A.G.; Pereira Jr., H.A.; Pezuelo, J.B.; Quaglio, R.B.; Roberto, P.G.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki, H.E.  
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira, J.F.; Silvestri, M.L.Z.; Siqueira, W.J.; de Souza, A.A.; de Souza, A.P.; Terenzi, M.F.; Trufil, D.; Tsai, S.M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Zago, M.A.; Zatz, M.; Meidanis, J.; Setubal, J.C.  
 A:Reference number: A59328  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: XF1562

Query Match 100.0%; Score 21; DB 2; Length 88;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGD 4  
 ||||  
 Db 65 RGD 68

## RESULT 15

168553

cell surface glycoprotein - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 04-Oct-1996 #sequence\_revision 04-Oct-1996 #text\_change 23-Jul-1999

C:Accession: 168553

R:Horn, G.T.; Bugawan, T.L.; Long, C.M.; Manos, M.M.; Erlich, H.A.

Hum. Immunol. 21, 249-263, 1998

A:Title: Sequence analysis of HLA class II genes from insulin-dependent diabetic individuals.

A:Reference number: 154290; MID:86227495; PMID:3372263

A:Accession: 168553

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-89 &lt;RES&gt;

A:Cross-references: GB:M35000; NID:g291960; PIRN:AA35774.1; PID:g553265

C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

C:Keywords: glycoprotein

Query Match 100.0%; Score 21; DB 2; Length 89;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGD 4  
 ||||  
 Db 44 RGD 47

Search completed: February 11, 2004, 14:56:56

OM protein - protein search, using sw model

Runon: February 11, 2004, 14:36:52 ; Search time 1.67742 seconds

(without alignments)  
112.141 Million cell updates/sec

Title: US-10-050-611-1

Perfect score: 21

Sequence: 1 RCDA 4

Scoring table: ELOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.1

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	21	100.0	39 1 DECO_MACDE	P17350 macrordella
2	21	100.0	52 1 ORNC_PLAOR	P25512 placordella
3	21	100.0	74 1 R38B_SCHPO	Q09900 schizosach
4	21	100.0	76 1 MTRB_BACPU	P48064 bacillus pu
5	21	100.0	80 1 C551_CHRVI	P80549 chromatium
6	21	100.0	97 1 RLZ1_PYRAB	Q94261 pyrococcus
7	21	100.0	97 1 RLZ1_PYRHO	Q74001 pyrococcus
8	21	100.0	98 1 ULI9_HCVNA	P16723 human cytom
9	21	100.0	113 1 AP61_HUMAN	Q15772 homo sapien
10	21	100.0	113 1 AP61_MOUSE	Q62407 mus musculu
11	21	100.0	113 1 AP61_RAT	Q62407 rattus norv
12	21	100.0	116 1 RL17_HELPJ	Q92166 helicobacte
13	21	100.0	116 1 RL17_HELPY	P56042 helicobacte
14	21	100.0	124 1 RL17_MYCPN	Q95647 mycoplasma
15	21	100.0	124 1 RSEB_HALNI	Q94969 halobacteri
16	21	100.0	131 1 RL17_THEMA	Q94111 thermotoga
17	21	100.0	133 1 GEPE_BACSU	Q06717 bacillus su

ALIGNMENTS

RESULT 1	ID	DECO_MACDE	STANDARD:	PRT:	39 AA.
AC	P17350;				
DT	01-AUG-1990 (Rel. 15, Created)				
DT	01-AUG-1990 (Rel. 15, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Decorin.				
OS	Macrordella decora (North American leech).				
OC	Eukaryota; Metazoa; Annelida; Clitellata; Hirudinea; Hirudinea;				
OC	Arymchordellidae; Hirudiniiformes; Hirudindae; Macrordella.				
OX	NCBI_Taxid=6405;				
RN	[1]				
RP	SEQUENCE.				
RX	MEDLINE=90277628; PubMed=2351655;				
RA	Seymour J.T., Henzel W.J., Nevins B., Stults J.T., Lazarus R.A.;				
RT	"Decorin. A potent glycoprotein 11b-IIIa antagonist and platelet				
RT	aggregation inhibitor from the leech Macrordella decora.";				
RL	J. Biol. Chem. 265:10143-10147(1990).				
RN	[2]				
RP	STRUCTURE BY NMR.				
RX	MEDLINE=94278302; PubMed=8009227;				
RA	Kretzel A.M., Wagner G., Seymour-Ulmer J., Lazarus R.A.;				
RT	"Structure of the RGD protein decorin: conserved motif and distinct				

18	21	100.0	140 1 COBB_RAT	P55314 rattus norv
19	21	100.0	141 1 NIKR_METJA	Q57969 methanococ
20	21	100.0	143 1 IR09_HCVNA	P16807 human cytom
21	21	100.0	149 1 DUT_CORGL	Q8npa9 corynebacte
22	21	100.0	150 1 FLA2_MERVO	Q06640 methanococ
23	21	100.0	150 1 MOKE_HAELN	P45308 haemophilus
24	21	100.0	151 1 CP2B_DROME	Q9n1p6 drosophila
25	21	100.0	155 1 RR7_CUSEU	P46292 cuscutea eur
26	21	100.0	157 1 Y510_VIBCH	Q9kxk8 vibrio chol
27	21	100.0	164 1 RL15_HALVA	P12737 halorcula
28	21	100.0	168 1 TPX_CHITE	Q8ked5 chlorobium
29	21	100.0	172 1 LBD4_ARATH	Q9ehs9 arabidopsis
30	21	100.0	177 1 RLE_HALVA	P14153 halorcula
31	21	100.0	179 1 YF36_PSEAE	Q913h7 pseudomonas
32	21	100.0	181 1 YG66_STROO	Q94266 streptomyc
33	21	100.0	185 1 RRF_BUCAI	P57328 buchnera ap
34	21	100.0	186 1 YCE7_DROME	Q97067 drosophila
35	21	100.0	190 1 Y2H5_STROO	P35925 streptomyc
36	21	100.0	192 1 TERD_MLSCP	P18781 alecigenes
37	21	100.0	197 1 HANU_PSEAE	Q916a8 pseudomonas
38	21	100.0	201 1 EFA4_HUMAN	P52798 homo sapien
39	21	100.0	201 1 SODE_ONCVO	Q07449 onchocerca
40	21	100.0	202 1 B3G1_MOUSE	Q9cm73 m galactosy
41	21	100.0	203 1 IDI_MXCTU	P72002 mycobacteri
42	21	100.0	206 1 EFA4_MOUSE	Q08542 mus musculu
43	21	100.0	206 1 YWAB_BACSU	P50619 bacillus su
44	21	100.0	212 1 RL17_HUMAN	Q94067 homo sapien
45	21	100.0	214 1 RADQ_VIBVU	Q8dd40 vibrio vuln

RT function in leech proteins that affect blood clotting."  
 RL Science 264:1944-1947(1994).  
 CC -1- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS  
 CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. MAY PREVENT BLOOD FROM  
 CC CLOTTING DURING EITHER FEEDING AND/OR STORAGE OF INGESTED BLOOD.  
 CC -1- SIMILARITY: HIGH, TO P. ORNATA ORNATINS.  
 CC PIR; A36453; A36453.  
 DR -1- SIMILARITY: SOME, TO THE DISINTEGRIN FAMILY.  
 DR PDB; IDEC; 31-AUG-94.  
 KW Blood coagulation; Platelet; Cell adhesion; 3D-structure.  
 FT DOMAIN 27 38 HIGH AFFINITY BINDING DOMAIN (POTENTIAL).  
 FT SITE 31 33 CELL ATTACHMENT SITE.  
 FT VARIANT 1 3 MISSING (IN N-3 ISOFORM).  
 FT STRAND 6 6  
 FT STRAND 15 16  
 FT STRAND 21 22  
 FT TURN 24 25  
 FT STRAND 27 28  
 FT STRAND 37 39  
 SQ SEQUENCE 39 AA; 4384 MW; 3A3B35756FB70D36 CRC64;  
 Query Match 100.0%; Score 21; DB 1; Length 39;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDA 4  
 DB 31 RGDA 34  
 RESULT 2  
 ID CRNC\_PLAOR STANDARD; PRT; 52 AA.  
 AC P25512;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ornatin C.  
 OS Placodonta ornata (Turtle leech).  
 OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinea; Hirudinea;  
 OC Rhyachobdellida; Glossiphoniidae; Placodonta.  
 NC NCB1\_TaxID=6415;  
 NM [1]  
 RP SEQUENCE.  
 RA MEDLINE=92111479; PubMed=1765068;  
 RA Mazur P., Henzel W.J., Seymour J.L., Lazarus R.A.;  
 RA Ornatin: potent glycoprotein IIB-IIIA antagonists and platelet  
 RA aggregation inhibitors from the leech Placodonta ornata."  
 RL Eur. J. Biochem. 202:1073-1082(1991).  
 CC -1- FUNCTION: POTENT INHIBITOR OF FIBRINOGEN INTERACTION WITH PLATELET  
 CC RECEPTORS EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. MAY PREVENT  
 CC BLOOD FROM CLOTTING DURING EITHER FEEDING AND/OR STORAGE OF  
 CC INGESTED BLOOD.  
 CC -1- SIMILARITY: BELONGS TO THE ORNATIN FAMILY.  
 CC PIR; S19623; S19623.  
 DR InterPro; IPR002463; Ornatin.  
 DR Pfam; PF02088; Ornatin; 1.

DR ProDom; PD012062; Ornatin; 1.  
 KW Blood coagulation; Platelet; Cell adhesion.  
 FT SITE 42 44 CELL ATTACHMENT SITE.  
 SQ SEQUENCE 52 AA; 5845 MW; BA53CA7408EF4F09 CRC64;  
 Query Match 100.0%; Score 21; DB 1; Length 52;  
 Best Local Similarity 100.0%; Pred. No. 66;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDA 4  
 DB 42 RGDA 45  
 RESULT 3  
 ID R38B\_SCHPO STANDARD; PRT; 74 AA.  
 AC 009900;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE 60S ribosomal protein L38-2.  
 GN RPL38B OR RPL38 OR SPAC30D11.12.  
 OS Schizosaccharomyces pombe (Fission Yeast).  
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 CC Schizosaccharomycetaceae; Schizosaccharomycetaceae;  
 CC Schizosaccharomycetes.  
 NC NCB1\_TaxID=4896;  
 NM [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21648401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Boman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Ungels K.,  
 RA James K., Jones L., Jones M., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Voiclaert G., Aert R., Robben J., Grymonprez B.,  
 RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Koestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Heinrich H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambut R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motlier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Huzar S.M.,  
 RA Lucas M., Rocher M., Gallardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Amatrone J., Forsburg S.L.,  
 RA Cerritoli L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nare P.,  
 RA "The genome sequence of Schizosaccharomyces pombe."

RL Nature 415:871-880(2002).  
 CC -1- MISCELLANEOUS: There are two genes for l38 in s.pombe.  
 CC -1- SIMILARITY: BELONGS TO THE L38 FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: Z67961; CAA91898.1; -.  
 CC PIR: S62570; S62570.  
 CC GeneDB: Spombe; SPAC30D11.12f -.  
 CC InterPro: IPR002675; Ribosomal\_L38e.  
 CC Pfam: PF01781; Ribosomal\_L38e; 1.  
 CC ProDom: PD010361; Ribosomal\_L38e; 1.  
 CC Ribosomal protein, Kallitene family.  
 CC SQUENCE 74 AA; 8339 MW; C90D6594DFCB11D3 CRC64;  
 SQ  
 Query Match 100.0%; Score 21; DB 1; Length 74;  
 Best Local Similarity 100.0%; Pred. No. 94;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDA 4  
 DB 17 RGDA 20  
 RESULT 4  
 MTRB\_BACPU STANDARD; PRT; 76 AA.  
 ID MTRB\_BACPU  
 AC P48064;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Transcription attenuation protein mtrb (tryptophan RNA-binding  
 DE attenuator protein) (Ttp RNA-binding attenuation protein) (TRAP).  
 GN MTRB.  
 OS Bacillus pumilus (Bacillus mesentericus).  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1408;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95138053; PubMed=7836324;  
 RA Hofman R.J., Golinick P.;  
 RT "The mtrb gene of Bacillus pumilus encodes a protein with sequence  
 RT and functional homology to the ttp RNA-binding attenuation protein  
 RT (TRAP) of Bacillus subtilis.";  
 RL J. Bacteriol. 177:839-842(1995).  
 CC -1- FUNCTION: REQUIRED FOR TRANSCRIPTION ATTENUATION CONTROL IN THE  
 CC NUCLEOTIDE SEQUENCE IN THE TRP LEADER TRANSCRIPT CAUSING A 10 BASES  
 CC TRANSCRIPTION TERMINATION. BINDS THE LEADER RNA ONLY IN PRESENCE  
 CC OF L-TRYPTOPHAN.  
 CC -1- SUBUNIT: OLIGOMER OF 11 IDENTICAL SUBUNITS ARRANGED IN DOUGHNUT-

CC LIKE STRUCTURE (BY SIMILARITY).  
 CC -1- SIMILARITY: WITH REGA FROM PHAGE T4.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: I37879; AA67544.1; -.  
 CC PIR: I39905; I39905.  
 CC HSSP: Q9X606; 10AW.  
 CC InterPro: IPR000824; TRBP.  
 CC Pfam: PF02081; TRBP; 1.  
 CC PRINTS: PR00687; TRPNNAP.  
 CC ProDom: PD027918; TRBP; 1.  
 CC Transcription regulation; RNA-binding.  
 CC SQUENCE 76 AA; 8301 MW; 22184B2351DA151D CRC64;  
 SQ  
 Query Match 100.0%; Score 21; DB 1; Length 76;  
 Best Local Similarity 100.0%; Pred. No. 97;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDA 4  
 DB 58 RGDA 61  
 RESULT 5  
 C551\_CHRYI STANDARD; PRT; 80 AA.  
 ID C551\_CHRYI  
 AC P80549;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Cytochrome c-551 (C551).  
 DE Chromatium vinosum.  
 OS Chromatium vinosum.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales;  
 OC Chromatiaceae; Allochrocatium.  
 OX NCBI\_TaxID=1049;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=D / ATCC 17899 / DSM 180;  
 RX MEDLINE=96193682; PubMed=8612646;  
 RA Samyn B., de Smet L., van Driessche G., Meyer T.E., Bartsch R.G.,  
 RA Cusanovitch M.A., van Beeumen J.J.;  
 RT "A high-potential soluble cytochrome c-551 from the purple  
 RT phototrophic bacterium Chromatium vinosum is homologous to cytochrome  
 RT c8 from denitrifying pseudomonads.";  
 RL Eur. J. Biochem. 236:689-696(1996).  
 CC -1- FUNCTION: MONOHEME CYTOCHROME.  
 CC PIR: S68677; S68677.  
 CC HSSP: P93339; 1X56.  
 CC InterPro: IPR003086; Cyt\_C1.  
 CC InterPro: IPR002324; Cyt\_C1D.

DR InterPro: IPR000345; Cytochrome bind.  
 DR Pfam: PF00034; cytochrome c 1.  
 DR PRINTS: PR000606; CYTOCHROME\_C.1.  
 DR PROSITE: PS00190; CYTOCHROME\_C.1.  
 KW Electron transport; Heme.  
 FT BINDING 10 10 HEME (COVALENT).  
 FT BINDING 13 13 HEME (COVALENT).  
 FT METAL 14 14 IRON (HEME AXIAL LIGAND).  
 FT METAL 59 59 IRON (HEME AXIAL LIGAND).  
 SQ SEQUENCE 80 AA; 6224 MW; EBD30A2815D07F93 CRC64;  
 Query Match 100.0%; Score 21; DB 1; Length 80;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDA 4  
 DB 33 RGDA 36  
 RESULT 6  
 RL21\_PYRAB STANDARD; PRT; 97 AA.  
 ID OSU2B1;  
 AC 09U2B1;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE 50S ribosomal protein L21e.  
 GN RPL21E OR PYRAB11050 OR PAB0731.  
 OS Pyrococcus abyssi.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 CC Pyrococcus.  
 OX NCBI\_TaxID=29292;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GB3 / Orsay;  
 RX PubMed=12622808;  
 RA Cohen G.N., Barbe V., Flament D., Galperin M., Hellig R., Lecompte O.,  
 RA Pech O., Pileur D., Querellou J., Ripp R., Thierry J.-C.,  
 RA Van der Oost J., Weissenbach J., Zivanovic Y., Fosterer P.;  
 RT "An integrated analysis of the genome of the hyperthermophilic  
 archaeon Pyrococcus abyssi.";  
 RL Mol. Microbiol. 47:1495-1512(2003).  
 CC -1- SIMILARITY: BELONGS TO THE L21E FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AJ248286; CAB50016.1; -.  
 DR PIR: C75089; C75089.  
 DR HAMAP: MF\_00369; -; 1.  
 DR InterPro: IPR001147; Ribosomal\_L21e.

DR Pfam: PF01157; Ribosomal\_L21e; 1.  
 DR PROSITE: PS01171; RIBOSOMAL\_L21E; 1.  
 KW Ribosomal protein; Complete proteome.  
 SQ SEQUENCE 97 AA; 11378 MW; 6CER3A2DB6A61E40 CRC64;  
 Query Match 100.0%; Score 21; DB 1; Length 97;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDA 4  
 DB 69 RGDA 72  
 RESULT 7  
 RL21\_PYRHO STANDARD; PRT; 97 AA.  
 ID AC 074001;  
 AC 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE 50S ribosomal protein L21e.  
 GN RPL21E OR PH127.1 OR PH5032.  
 OS Pyrococcus horikoshii.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 CC Pyrococcus.  
 OX NCBI\_TaxID=53953;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OT3;  
 RX MEDLINE=98344137; PubMed=9679194;  
 RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hirao Y.,  
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Haseyama A., Nagai Y.,  
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Onitoku Y.,  
 RA Funahashi T., Tanaka T., Kudo Y., Yamazaki J., Kusuda N., Oguchi A.,  
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
 RA Masuchi Y., Shizuya H., Kikuchi H.;  
 RT "Complete sequence and gene organization of the genome of a hyper-  
 RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
 RL DNA Res. 5:55-76(1998).  
 CC -1- SIMILARITY: BELONGS TO THE L21E FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF000005; BAA30227.1; -.  
 DR PIR: A71054; A71054.  
 DR HAMAP: MF\_00069; -; 1.  
 DR InterPro: IPR001147; Ribosomal\_L21e.  
 DR Pfam: PF01157; Ribosomal\_L21e; 1.  
 DR PROSITE: PS01171; RIBOSOMAL\_L21E; 1.  
 KW Ribosomal protein; Complete proteome.

SQ SEQUENCE 97 AA; 11376 MW; 6D5D229DBFBE0E51 CRCE4;  
 Query Match 100.0%; Score 21; DB 1; Length 97;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4  
 ||||  
 DB 69 RGDA 72

RESULT 8  
 ID U119\_HQCV A STANDARD; PRT; 98 AA.  
 AC P16723;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DE 01-FEB-1991 (Rel. 17, Last annotation update)  
 DE Hypothetical protein U119.  
 GN U119.  
 OS Human cytomegalovirus (strain AD169).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Betaherpesvirinae; Cytomegalovirus.  
 OX NCBI\_TaxID=10360;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88094735; PubMed=2827039;  
 RA Beck S., Barrell B.G.;  
 RT "Human cytomegalovirus encodes a glycoprotein homologous to MHC  
 RT class-I antigens";  
 RL Nature 331:269-272(1988).  
 RN [2]  
 RP COMPLETE GENOME.  
 RX MEDLINE=90269039; PubMed=2161319;  
 RA Chee M.S., Bankier A.T., Beck S., Bohm R., Brown C.M., Cerny R.,  
 RA Horsnell T., Hutchison C.A. III, Kourzides T., Martignetti J.A.,  
 RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;  
 RT "Analysis of the protein-coding content of the sequence of human  
 RT cytomegalovirus strain AD169";  
 RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; Y00293; -; NOT ANNOTATED\_CDS.  
 DR EMBL; X17403; CA35418.1; -.  
 DR PIR; S01566; S01566.  
 KW Hypothetical protein.  
 SQ SEQUENCE 98 AA; 11280 MW; 7E9A7405611E3F2B CRCE4;  
 Query Match 100.0%; Score 21; DB 1; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4  
 ||||  
 DB 95 RGDA 98

RESULT 9  
 ID AP01\_HUMAN STANDARD; PRT; 113 AA.  
 AC Q15772;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Aortic preferentially expressed protein 1 (APEG-1).  
 GN APEG1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=966291890; PubMed=8663449;  
 RA Hsieh C.-M., Yoshizumi M., Endoge W.O., Kho C.-J., Jain M.K.,  
 RA Kashiki S., de Los Santos R., Lee W.-S., Perrella M.A., Lee M.-E.;  
 RT "APEG-1, a novel gene preferentially expressed in aortic smooth muscle  
 RT cells, is down-regulated by vascular injury";  
 RL J. Biol. Chem. 271:17354-17359(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.K., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz J.E.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Botak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Wooley K.C., Hale S., Garcia A.M., Gay L.J., Huzyk S.W.,  
 RA Villalón D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kertman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Biskreley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schultz J., Myers R.W.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Skalska D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Meara M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).  
 CC -1- FUNCTION: MAY HAVE A ROLE IN REGULATING THE GROWTH AND  
 CC DIFFERENTIATION OF ARTERIAL SMOOTH MUSCLE CELLS.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN DIFFERENTIATED  
 CC ARTERIAL SMOOTH MUSCLE CELLS (ASMC).

CC -1- DEVELOPMENTAL STAGE: APPEARS TO BE EXPRESSED ONLY IN HIGHLY  
 CC DIFFERENTIATED ASMC IN NORMAL VESSEL WALLS AND DOWN-REGULATED IN  
 CC DIFFERENTIATED ASMC IN VIVO. IN RESPONSE TO VASCULAR INJURIES  
 CC ASMC DIFFERENTIATE AND CHANGE FROM A QUIESCENT AND CONTRACTILE  
 CC PHENOTYPE TO A PROLIFERATIVE AND SYNTHETIC PHENOTYPE. THIS  
 CC PROLIFERATION OF VASCULAR SMOOTH MUSCLE CELLS IS ONE OF THE MOST  
 CC PROMINENT FEATURES OF ARTERIOSCLEROSIS.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; U57099; AAC50599.1; -.  
 CC DR EMBL; BC006346; AAH06346.1; -.  
 CC DR HSSP; P56276; 1TLK.  
 CC DR GO; GO:0005634; C:nucleus; TAS.  
 CC DR GO; GO:0007517; P:muscle development; TAS.  
 CC DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.  
 CC DR InterPro; IPR007110; Ig-like.  
 CC DR InterPro; IPR003598; Ig\_c2.  
 CC DR InterPro; IPR003006; Ig\_MHC.  
 CC DR Pfam; PF00047; Ig\_1.  
 CC DR SMART; SM00408; Ig\_c2; 1.  
 CC DR PROSITE; PS00835; IG\_LIKE; 1.  
 CC KW Immunoglobulin domain; Nuclear protein.  
 CC FT DOMAIN 20 109 IG-LIKE.  
 CC SQ SEQUENCE 113 AA; 12692 MW; 04F367263A1397C5 CRC64;  
 CC  
 CC Query Match 100.0%; Score 21; DB 1; Length 113;  
 CC Best Local Similarity 100.0%; Pred.No. 1.4e+02;  
 CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 1 RGDA 4  
 CC DB 85 RGDA 88  
 CC  
 CC RESULT 10  
 CC AP01\_MOUSE  
 CC ID AP01\_MOUSE STANDARD; PRT; 113 AA.  
 CC AC Q62407;  
 CC DT 30-MAY-2000 (Rel. 39, Created)  
 CC DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 CC DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 CC DE Aortic preferentially expressed protein 1 (APEG-1).  
 CC GN APEG1.  
 CC OS Mus musculus (Mouse).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CC OX NCBI\_TaxID=10090;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6;  
 RX MEDLINE=96291890; PubMed=8663449;  
 RA Hsieh C.-M., Yoshizumi M., Endege W.O., Kho C.-J., Jain M.K.,  
 RA Kashi S., de Los Santos R., Lee W.-S., Petrella M.A., Lee M.-E.,  
 RT "APEG-1, a novel gene preferentially expressed in aortic smooth muscle  
 RT cells, is down-regulated by vascular injury,"  
 RL J. Biol. Chem. 271:17354-17359(1996).  
 CC -1- FUNCTION: MAY HAVE A ROLE IN REGULATING THE GROWTH AND  
 CC DIFFERENTIATION OF ARTERIAL SMOOTH MUSCLE CELLS.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN DIFFERENTIATED  
 CC ARTERIAL SMOOTH MUSCLE CELLS (ASMC).  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; U57098; AAC52666.1; -.  
 CC DR HSSP; P56276; 1TLK.  
 CC DR MED; MG1:109282; Apeg1.  
 CC DR InterPro; IPR007110; Ig-like.  
 CC DR InterPro; IPR003598; Ig\_c2.  
 CC DR InterPro; IPR003006; Ig\_MHC.  
 CC DR Pfam; PF00047; Ig\_1.  
 CC DR SMART; SM00408; Ig\_c2; 1.  
 CC DR PROSITE; PS00835; IG\_LIKE; 1.  
 CC KW Immunoglobulin domain; Nuclear protein.  
 CC FT DOMAIN 20 109 IG-LIKE.  
 CC SQ SEQUENCE 113 AA; 12665 MW; 5F320C5A41C3D870 CRC64;  
 CC  
 CC Query Match 100.0%; Score 21; DB 1; Length 113;  
 CC Best Local Similarity 100.0%; Pred.No. 1.4e+02;  
 CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 1 RGDA 4  
 CC DB 85 RGDA 88  
 CC  
 CC RESULT 11  
 CC AP01\_RAT  
 CC ID AP01\_RAT STANDARD; PRT; 113 AA.  
 CC AC Q63638;  
 CC DT 30-MAY-2000 (Rel. 39, Created)  
 CC DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 CC DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 CC DE Aortic preferentially expressed protein 1 (APEG-1).  
 CC GN APEG1.  
 CC OS Rattus norvegicus (Rat).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 CC OX NCBI\_TaxID=10116;  
 CC RP SEQUENCE FROM N.A.

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley;  
 RX MEDLINE=96231890; PubMed=8663449;  
 RA Hashin C.-M., Yoshizumi M., Endege W.O., Kho C.-J., Jain M.K.,  
 RA Kashiiki S., de Los Santos R., Lee W.-S., Perrella M.A., Lee M.-E.,  
 RT "APEG-1, a novel gene preferentially expressed in aortic smooth muscle  
 RT cells, is down-regulated by vascular injury."  
 RL J. Biol. Chem. 271:17354-17359(1996).  
 CC -|- FUNCTION: MAY HAVE A ROLE IN REGULATING THE GROWTH AND  
 CC DIFFERENTIATION OF ARTERIAL SMOOTH MUSCLE CELLS.  
 CC -|- SUBCELLULAR LOCATION: Nuclear.  
 CC -|- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN DIFFERENTIATED ARTERIAL  
 CC SMOOTH MUSCLE CELLS (ASMC) IN THE MEDIAL LAYER OF THE AORTA.  
 CC WEAKLY DETECTED IN BRAIN AND TESTIS AND TO A LESSER EXTENT IN  
 CC ORGANS RICH IN STRIATED MUSCLE OR VISCERAL SMOOTH MUSCLE.  
 CC -|- SIMILARITY: Contains 1 immunoglobulin-like domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U57097; AAC52667.1; -.  
 DR HSSP; P56276; 1TLK.  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003598; IG\_C2.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR Pfam; PF00047; 1g\_1.  
 DR SMART; SM00408; IgC2; 1.  
 DR PROSITE; PSS0835; IG\_LIKE; 1.  
 DR Immunoglobulin domain; Nuclear protein.  
 KW DOMAIN  
 FT 20 109 IG-LIKE.  
 SQ SEQUENCE 113 AA; 12668 MW; B213C366A759A363 CRC64;  
 QY Query Match 100.0%; Score 21; DB 1; Length 113;  
 DB Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDA 4  
 DB 85 RGDA 86  
 RESULT 12  
 ID RL17\_HELPY STANDARD; PRT; 116 AA.  
 AC 092JTG;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 50S ribosomal protein L17.  
 GN RPLQ OR HPI212.  
 OS Helicobacter pylori J99 (Campylobacter pylori J99).

OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;  
 OC Helicobacteraceae; Helicobacter.  
 OX NCBI\_TaxID=85963;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99120537; PubMed=9923682;  
 RA Alm R.A., Ling L.-S.L., Wolf D.T., King B.L., Brown E.D., Dolg P.C.,  
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,  
 RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,  
 RA Gibson R., Werberg D., Mills S.D., Jiang Q., Taylor D.E., Voyts G.F.,  
 RA Trust T.J.;  
 RT "Genomic sequence comparison of two unrelated isolates of the human  
 RT gastric pathogen Helicobacter pylori."  
 RL Nature 397:176-180(1999).  
 CC -|- SIMILARITY: BELONGS TO THE L17P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AB001547; AB006814.1; -.  
 DR PIR; D71832; D71832.  
 DR InterPro; IPR000456; Ribosomal\_L17.  
 DR Pfam; PF01196; Ribosomal\_L17; 1.  
 DR ProDom; PD004277; Ribosomal\_L17; 1.  
 DR TIGRfams; TIGR00059; L17; 1.  
 DR PROSITE; PS01167; RIBOSOMAL\_L17; 1.  
 KW Ribosomal protein; Complete proteome.  
 SQ SEQUENCE 116 AA; 13392 MW; EBC77780E2F2F3A1 CRC64;  
 QY Query Match 100.0%; Score 21; DB 1; Length 116;  
 DB Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RGDA 4  
 DB 104 RGDA 107  
 RESULT 13  
 ID RL17\_HELPY STANDARD; PRT; 116 AA.  
 AC P36042;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 50S ribosomal protein L17.  
 GN RPLQ OR HPI292.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;  
 OC Helicobacteraceae; Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]



OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
 OC Halobacteriaceae; Halobacterium.  
 OX NCBI\_TaxID=64091;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20504483; PubMed=1016950;  
 RA Ng W.V., Kennedy S.P., Kahliras G.G., Bergquist B., Pan M.,  
 RA Shukla H.D., Lasky S.R., Balliga N.S., Thorsson V., Spogna J.,  
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Weir R., Goo Y.A.,  
 RA Leithauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angermeier C.M., Dale H.,  
 RA Isernberger T.A., Peck K.F., Fornschröder M., Spudis J.L., Jung K.-H.,  
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
 RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;  
 RA "Genome sequence of Halobacterium species NRC-1.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
 CC -1- SIMILARITY: BELONGS TO THE SBE FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@ebi-sib.ch](mailto:license@ebi-sib.ch)).  
 CC -----  
 DR EMBL; AE005076; AAG19920.1; -.  
 DR PIR; D64319; D64319.  
 DR HAMAP; ME\_00029; -; 1.  
 DR InterPro; IPR001047; Ribosomal\_SBE.  
 DR Pfam; PF01201; Ribosomal\_SBE; 1.  
 DR ProDom; PD005658; Ribosomal\_SBE; 1.  
 DR TIGRFAMs; TIGR00307; Sbe; 1.  
 DR PROSITE; PS01193; RIBOSOMAL\_SBE; 1.  
 KW Ribosomal protein; Complete proteome.  
 SQ SEQUENCE 124 AA; 13515 MW; B7038CF79A83742B CRC64;

Query Match 100.0%; Score 21; DB 1; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4  
 ||||  
 DB 47 RGDA 50

Search completed: February 11, 2004, 14:54:03  
 Job time : 4.67742 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 14:47:57 ; Search time 6.83871 Seconds

(without alignments)  
 150.936 Million cell updates/sec

Title: US-10-050-611-1

Perfect score: 21

Sequence: 1 RGDA 4

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database : SPTREMBL\_23:\*
- 1: sp\_archaea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_phage:\*
  - 10: sp\_plant:\*
  - 11: sp\_protist:\*
  - 12: sp\_virus:\*
  - 13: sp\_vertebrate:\*
  - 14: sp\_unclassified:\*
  - 15: sp\_rvirs:\*
  - 16: sp\_bacteriopl:\*
  - 17: sp\_archesp:\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query Score	Match length	DB ID	Description
-----	-----	-----	-----	-----

1	21	100.0	31	5	Q8XEB8	Q8XEB8 caenorhabdi
2	21	100.0	45	16	Q9PGB6	Q9PGB6 xylella fas
3	21	100.0	48	2	Q9XDV3	Q9XDV3 erythroba
4	21	100.0	54	16	Q8RTH3	Q8RTH3 thermomere
5	21	100.0	55	10	Q8RZ21	Q8RZ21 zea mays (m
6	21	100.0	57	6	Q9NC41	Q9NC41 mactaca fasc
7	21	100.0	57	10	Q8RUD5	Q8RUD5 zea mays (m
8	21	100.0	57	10	Q8RUD4	Q8RUD4 zea mays (m
9	21	100.0	57	16	Q8R773	Q8R773 mycobacteri
10	21	100.0	58	12	Q8Q583	Q8Q583 chimpanzee
11	21	100.0	59	16	Q8EL57	Q8EL57 rhicobium l
12	21	100.0	64	16	Q8XJQ0	Q8XJQ0 ralteonia s
13	21	100.0	66	12	Q8JXZ2	Q8JXZ2 virus phich
14	21	100.0	68	5	Q8MNA5	Q8MNA5 dictyosteli
15	21	100.0	68	16	Q8UYK6	Q8UYK6 agrobacteri
16	21	100.0	69	16	Q8DDL7	Q8DDL7 vibrio vuln
17	21	100.0	70	12	Q8VAV0	Q8VAV0 white spot
18	21	100.0	70	16	Q8XTW3	Q8XTW3 ralteonia s
19	21	100.0	73	16	Q8Y128	Q8Y128 ralteonia s
20	21	100.0	75	16	Q8VJ45	Q8VJ45 mycobacteri
21	21	100.0	76	10	Q8GVK2	Q8GVK2 cryza sativ
22	21	100.0	77	6	Q29171	Q29171 sus scrofa
23	21	100.0	77	16	Q92X10	Q92X10 rhicobium m
24	21	100.0	79	16	Q8XG07	Q8XG07 escherichia
25	21	100.0	83	17	Q8TK40	Q8TK40 methanocarc
26	21	100.0	85	10	Q8W3B8	Q8W3B8 cryza sativ
27	21	100.0	88	16	Q9PD18	Q9PD18 xylella fas
28	21	100.0	88	17	Q8ZV78	Q8ZV78 pyrobaculum
29	21	100.0	89	5	Q95Y01	Q95Y01 caenorhabdi
30	21	100.0	89	7	Q29783	Q29783 homo sapien
31	21	100.0	89	16	Q8G801	Q8G801 bifidobacte
32	21	100.0	90	16	Q9PAM0	Q9PAM0 xylella fas
33	21	100.0	91	15	Q9DR41	Q9DR41 human immun
34	21	100.0	91	16	Q8PJH2	Q8PJH2 xanthomonas
35	21	100.0	92	9	Q9FZ75	Q9FZ75 pseudomonas
36	21	100.0	93	10	Q8S2D8	Q8S2D8 cryza sativ
37	21	100.0	93	16	Q8Z7W3	Q8Z7W3 salmonella
38	21	100.0	95	16	Q9PDS1	Q9PDS1 xylella fas
39	21	100.0	95	16	Q9PTE9	Q9PTE9 rhizobium l
40	21	100.0	96	16	Q9KE84	Q9KE84 bacillus ha
41	21	100.0	96	17	Q9HR67	Q9HR67 halobacteri
42	21	100.0	97	16	Q9HTA8	Q9HTA8 pseudomonas
43	21	100.0	99	2	Q8RM68	Q8RM68 bacteroides
44	21	100.0	100	12	Q98Z39	Q98Z39 molluscum c
45	21	100.0	100	12	Q8B9W5	Q8B9W5 influenza b

ALIGNMENTS

RESULT 1  
Q8XEB8 ID Q8XEB8 PRELIMINARY; PRT; 31 AA.  
AC Q8XEB8; 01-OCT-2002 (TRENBLrel. 22, Created)  
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)

DE Hypothetical protein K07A9.4.  
GN K07A9.4.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=9069613; PubMed=9851916;  
RA Waterston R.;  
RT "Genome sequence of the nematode C. elegans: a platform for  
RT investigating biology. The C. elegans Sequencing Consortium."  
RL Science 282:2012-2018(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX "The sequence of C. elegans cosmid K07A9."  
RA Davidson S., O'Neal D.;  
RT Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX Waterston R.;  
RA Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AF099924; JAM98005.1; -.  
DR WormPep; K07A9.4; CE31709.  
KW Hypothetical protein.  
SQ SEQUENCE 31 AA; 3720 MW; 147938913DC940ED CRC64;

Query Match 100.0%; Score 21; DB 5; Length 31;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4  
Db 2 RGDA 5  
RESULT 2  
Q9PGB6 ID Q9PGB6 PRELIMINARY; PRT; 45 AA.  
AC Q9PGB6; 01-OCT-2000 (TRENBLrel. 15, Created)  
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)  
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
DE Hypothetical protein Kf0386.  
GN Kf0386.  
OS Xylella fastidiosa.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
OC Xanthomonadaceae; Xylella.  
OX NCBI\_TaxID=2371;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=9abc;  
RX MEDLINE=20365717; PubMed=10910347;  
RA Simpson A.J.G., Reinach F.C., Artuda P., Abreu F.A., Acencio M.,

RA Alverenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,  
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,  
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Garrão D.M., Carer H.,  
 RA Coluto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,  
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,  
 RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
 RA Ho P.L., Hobeisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
 RA Lemos E.G.H., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
 RA Machado M.A., Madela A.M.B.N., Madela H.M.F., Martino C.L.,  
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
 RA Moon D.H., Nagel M.A., Nascimento A.L.T.O., Netto L.E.S.,  
 RA Nhari A., Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A., Jr., Pesquero J.B.,  
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
 RA de Rosa V.E., Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A., Jr.,  
 RA da Silveira J.F., Silvestri M.L.Z., Silveira W.J., de Souza A.A.,  
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,  
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
 RA Zago M.A., Zatz M., Weidman J., Secudal J.C.,  
 RA "The genome sequence of the plant pathogen *Xylella fastidiosa*,"  
 RL Nature 406:151-159(2000).  
 DR EMBL: AE003890; AAF83196.1; -  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 45 AA; 5163 MW; B58C9AEC9809C8A CRC64;

Query Match 100.0%; Score 21; DB 16; Length 45;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGDA 4  
 1111  
 19 RGDA 22

Db 19 RGDA 22

RESULT 3  
 ID Q9XDY3 PRELIMINARY; PRT; 48 AA.  
 AC Q9XDY3:  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE ORF Q.  
 OS Erythrobacter sp. M81C3960.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;  
 CC Sphingomonadaceae; Erythrobacter.  
 OK NCBI\_TaxID=94771;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=M81C3960;  
 RA Hamada T.;  
 "Nucleotide sequences of genes coding for photosynthetic reaction

RI centers and light-harvesting proteins of Erythrobacter litoralis and  
 RI related aerobic photosynthetic bacteria,"  
 RI Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB027515; BAA78669.1; -  
 DR InterPro; IPR006089; Acyl-CoA.dh.  
 DR PROSITE; PS00073; ACIL\_COA\_Dh\_2; 1.  
 SQ SEQUENCE 48 AA; 4980 MW; D63EAD05EA8079B CRC64;

Query Match 100.0%; Score 21; DB 2; Length 48;  
 Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGDA 4  
 1111  
 27 RGDA 30

Db 27 RGDA 30

RESULT 4  
 ID Q8R7H3 PRELIMINARY; PRT; 54 AA.  
 AC Q8R7H3:  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Hypothetical protein TTE2436.  
 GN TTE2436.  
 OS Thermoaerobacter tengcongensis.  
 OC Bacteria; Firmicutes; Clostridia; Thermoaerobacteriales;  
 CC Thermoaerobacteriaceae; Thermoaerobacter.  
 OK NCBI\_TaxID=119072;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=M84 / JCM 111007;  
 RX MEDLINE=21992816; PubMed=11997336;  
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;  
 RT "A complete sequence of *T. tengcongensis* genome,"  
 RL Genome Res. 12:689-700(2002).  
 DR EMBL: AE019183; AM25571.1; -  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 54 AA; 6252 MW; 0A9C818C07DD905B CRC64;

Query Match 100.0%; Score 21; DB 16; Length 54;  
 Best Local Similarity 100.0%; Pred. No. 5.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGDA 4  
 1111  
 32 RGDA 35

Db 32 RGDA 35

RESULT 5  
 ID Q8RUZ1 PRELIMINARY; PRT; 55 AA.  
 AC Q8RUZ1:  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)  
 DE Acetyl-CoA C-acyltransferase-like protein (Fragment).  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.  
 OC NCBI\_TaxID=4577;  
 RX [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=Various strains;  
 RA Ching A.S., Caldwell K.S., Jung M., Dolan M., Smith O.S., Tingey S.,  
 RA Morgante M., Rafalski J.A.;  
 RT "SNP frequency, haplotype structure and linkage disequilibrium in  
 RT elite maize inbred lines."  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF498463; AAM14479.1; -;  
 DR EMBL; AF498469; AAM14485.1; -;  
 DR EMBL; AF498472; AAM14488.1; -;  
 DR EMBL; AF498477; AAM14493.1; -;  
 DR EMBL; AF498482; AAM14498.1; -;  
 DR EMBL; AF498485; AAM14501.1; -;  
 DR EMBL; AF498486; AAM14502.1; -;  
 DR InterPro: IPR002155; Thiolase.  
 DR Pfam: PF02803; thiolase\_C1.  
 DR PROSITE; PS00099; THIO\_LAS\_3; 1.  
 DR Acyltransferase; Transferase.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 55 AA; 5959 MW; 5C09DAC7224451D0 CRC64;  
 QY 1 RGDA 4  
 DB 31 RGDA 34  
 RESULT 6  
 Q9N041 PRELIMINARY; PRT; 57 AA.  
 AC Q9N041;  
 DT 01-OCT-2000 (TEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)  
 DE Unnamed protein product.  
 OS Macaca fascicularis (Crested macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecoidea; Macaca.  
 OC NCBI\_TaxID=9541;  
 RX [1]  
 RN SEQUENCE FROM N.A.  
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,  
 RA Suzuki Y., Sugano S., Hashimoto K.;  
 RT "Isolation of full-length cDNA clones from macaque brain cDNA

RI libraries."  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB046091; BAB01673.1; -;  
 SQ SEQUENCE 57 AA; 6250 MW; 300DE0C64A4897A9 CRC64;  
 QY 1 RGDA 4  
 DB 10 RGDA 13  
 RESULT 7  
 Q9RUD5 PRELIMINARY; PRT; 57 AA.  
 AC Q9RUD5;  
 DT 01-JUN-2002 (TEMBLrel. 21, Created)  
 DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)  
 DE Acetyl-CoA C-acyltransferase-like protein (Fragment).  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.  
 OC NCBI\_TaxID=4577;  
 RX [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=Various strains;  
 RA Ching A.S., Caldwell K.S., Jung M., Dolan M., Smith O.S., Tingey S.,  
 RA Morgante M., Rafalski J.A.;  
 RT "SNP frequency, haplotype structure and linkage disequilibrium in  
 RT elite maize inbred lines."  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF498457; AAM14473.1; -;  
 DR EMBL; AF498458; AAM14474.1; -;  
 DR EMBL; AF498459; AAM14475.1; -;  
 DR EMBL; AF498460; AAM14476.1; -;  
 DR EMBL; AF498461; AAM14477.1; -;  
 DR EMBL; AF498462; AAM14478.1; -;  
 DR EMBL; AF498464; AAM14480.1; -;  
 DR EMBL; AF498465; AAM14481.1; -;  
 DR EMBL; AF498466; AAM14482.1; -;  
 DR EMBL; AF498467; AAM14483.1; -;  
 DR EMBL; AF498468; AAM14484.1; -;  
 DR EMBL; AF498470; AAM14486.1; -;  
 DR EMBL; AF498471; AAM14487.1; -;  
 DR EMBL; AF498473; AAM14489.1; -;  
 DR EMBL; AF498475; AAM14491.1; -;  
 DR EMBL; AF498478; AAM14494.1; -;  
 DR EMBL; AF498480; AAM14496.1; -;  
 DR EMBL; AF498481; AAM14497.1; -;  
 DR EMBL; AF498483; AAM14499.1; -;  
 DR EMBL; AF498484; AAM14500.1; -;  
 DR EMBL; AF498487; AAM14503.1; -;  
 DR InterPro: IPR002155; Thiolase.

DR Pfam; PF02803; thiolase\_C; 1.  
 DR PROSITE; PS00099; THIOLEASE\_3; 1.  
 KW Acyltransferase; Transferase.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 57 AA; 6203 MW; DC4596C27A4451A8 CRC64;  
 Query Match 100.0%; Score 21; DB 10; Length 57;  
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGD4 4  
 |||||  
 DB 33 RGD4 36

RESULT 8  
 Q8RUD4 PRELIMINARY; PRT; 57 AA.  
 AC Q8RUD4;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Acetyl-CoA C-acyltransferase-like protein (Fragment).  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.  
 OX NCBI\_TaxID=4577;  
 RX [1]  
 KW Acyltransferase; Transferase.  
 RP STRAIN:cv. ITANA; cv. D71-4HT, and cv. H60;  
 RA Ching A.S., Caldwell K.S., Chung M., Dolan M., Smith O.S., Tingey S.,  
 RA Morgante M., Rafalski J.A.;  
 RT "SNP frequency, haplotype structure and linkage disequilibrium in  
 elite maize inbred lines."  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF498474; AAM4490.1; -;  
 DR EMBL; AF498476; AAM4492.1; -;  
 DR EMBL; AF498479; AAM4493.1; -;  
 DR InterPro; IPR002153; Thiolase.  
 DR Pfam; PF02803; thiolase\_C; 1.  
 DR PROSITE; PS00099; THIOLEASE\_3; 1.  
 KW Acyltransferase; Transferase.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 57 AA; 6185 MW; DC4596C76E4451A8 CRC64;  
 Query Match 100.0%; Score 21; DB 10; Length 57;  
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGD4 4  
 |||||  
 DB 33 RGD4 36

RESULT 9  
 006773

ID 006773 PRELIMINARY; PRT; 57 AA.  
 AC 006773;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical protein Rv0666.  
 GN Rv0666 OR MTC1376.10C.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinomycetales;  
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RX [1]  
 KW complete genome sequence.  
 RP STRAIN:H37Rv;  
 RC MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eigler K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Felkner T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sultson J.E., Taylor K., Whitehead S., Barrall B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 complete genome sequence."  
 RL Nature 393:537-544(1998).  
 DR EMBL; Z95972; CAB09391.1; -;  
 DR Tuberculat; Rv0666; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 57 AA; 5849 MW; 6285645BD7D0F2E CRC64;  
 Query Match 100.0%; Score 21; DB 16; Length 57;  
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGD4 4  
 |||||  
 DB 24 RGD4 27

RESULT 10  
 Q8Q583 PRELIMINARY; PRT; 58 AA.  
 AC Q8Q583;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE U12.  
 OS Chimpanzee cytomegalovirus.  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Betaherpesvirinae; Cytomegalovirus.  
 OX NCBI\_TaxID=18763;  
 RX [1]  
 KW complete genome sequence.  
 RP STRAIN:FROM N.A.  
 RA Davidson A.J., Akter P., Dolan A., Wright K.N., Addison C.,  
 RA Alencor D.J., Hayward G.S., McGeoch D.J.;  
 RT "The human cytomegalovirus genome revisited."

RU Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF480884; AA00694.1; -  
SQ SEQUENCE 58 AA; 6789 MW; 27400659BBD2BAD7 CRC64;

Query Match 100.0%; Score 21; DB 12; Length 58;  
Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGDA 4  
||||  
DB 2 RGDA 5

RESULT 11

Q98LS7 PRELIMINARY; PRT; 59 AA.  
AC Q98LS7;  
DT 01-OCT-2001 (TREMBLrel. 18, Created)  
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Hypothetical protein ms10897.  
GN MS10897.  
OS Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Phyllobacteriaceae; Mesorhizobium.  
OX NCBI\_TaxID=381;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MAFF303099;  
RX MEDLINE=21082930; PubMed=11214968;  
RA Keneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
RA Watanabe A., Ideawa K., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsumo A.,  
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,  
RA Takeuchi C., Yamada M., Tabata S.;  
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
R Mesorhizobium loti.";  
RL DNA Res. 7:331-338(2000).  
DR EMBL: AP002996; BAB48386.1; -  
KR Hypothetical protein; Complete proteome.  
SQ SEQUENCE 59 AA; 6059 MW; 4EE77EF3940B6633 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 59;  
Best Local Similarity 100.0%; Pred. No. 5.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGDA 4  
||||  
DB 36 RGDA 39

RESULT 12

Q8XY00 PRELIMINARY; PRT; 64 AA.  
AC Q8XY00;  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Hypothetical protein RSC1708.  
GN RSC1708 OR R502894.  
OS Ralstonia solanacearum (Pseudomonas solanacearum).  
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
OC Ralstoniaceae; Ralstonia.  
OX NCBI\_TaxID=303;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GM11000;  
RX MEDLINE=21691879; PubMed=11923852;  
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
RA Ariat M., Billault A., Brotlier P., Canus J.C., Catolico L.,  
RA Chandler M., Cholme N., Claudel-Renard C., Cunnac S., Demange N.,  
RA Gaspin C., Lave M., Moisan A., Robert C., Saurin W., Schlex T.,  
RA Siguer P., Thebaud P., Whalen M., Wincker P., Levy M.,  
RA Weisenbach J., Boucher C.A.;  
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";  
RL Nature 415:487-502(2002).  
DR EMBL: AL646066; CAD15410.1; -  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 64 AA; 7210 MW; F53F8AF55609609 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 64;  
Best Local Similarity 100.0%; Pred. No. 6.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGDA 4  
||||  
DB 60 RGDA 63

RESULT 13

Q8JKZ2 PRELIMINARY; PRT; 66 AA.  
AC Q8JKZ2;  
DT 01-OCT-2002 (TREMBLrel. 22, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE Hypothetical protein.  
OS Virus phich1.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.  
OX NCBI\_TaxID=114777;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20177831; PubMed=10712697;  
RA Baranyi U., Klein R., Lubitz W., Kruger D.H., Witte A.;  
RT "The archaeal halophilic virus-encoded Dam-like methyltransferase M.  
phich1-I methylates adenine residues and complements dam mutants in  
the low salt environment of Escherichia coli.";  
RL Mol. Microbiol. 35:1168-1179(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20497008; PubMed=11040128;  
RA Klein R., Greineder B., Baranyi U., Witte A.;  
RT "The structural protein E of the archaeal virus phich1: evidence for  
processing in Nautila magadii during virus maturation.";

RL Virology 276:376-387 (2000).  
 (3)  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=22136043; PubMed=12139629;  
 RA Klein R., Baranyi U., Roessler N., Greineder B., Scholz H., Witte A.;  
 RT "Nattalba magadii virus phiCh1: first complete nucleotide sequence  
 RT archaean.";  
 RT Mol. Microbiol. 45:851-863 (2002).  
 RN (4)  
 RP SEQUENCE FROM N.A.  
 RA Klein R., Baranyi U., Roessler N., Greineder B., Scholz H.;  
 RT "Sequence analysis of the temperate virus phiCh1 infecting the  
 RT haloalkaliphilic archaean Nattalba magadii.";  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF440695; AAK88738.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 66 AA; 6695 MW; 36EA1246C5F281A6 CRC64;  
 SO  
 Query Match 100.0%; Score 21; DB 12; Length 66;  
 Best Local Similarity 100.0%; Pred. No. 6.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 RGDA 4  
 ||||  
 Db 20 RGDA 23  
 RESULT 14  
 OBMN5 PRELIMINARY; PRT; 68 AA.  
 ID OBMN5  
 AC OBMN5;  
 DT 01-OCT-2002 (TEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TEMBLrel. 22, Last annotation update)  
 DE Hypothetical protein.  
 OS Dictyostelium discoideum (Slime mold).  
 CC Eukaryota; Mycetozoa; Dictyostelids; Dictyostelium.  
 OX NCBI\_TaxID=44689;  
 RX (1)  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=AX4;  
 RA Gloeckner G., Eichinger L., Szafranski K., Pachbat J., Dear P.,  
 RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,  
 RA Tungel B., Cox B., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;  
 RT "Sequence and Analysis of Chromosome 2 of Dictyostelium.";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC117076; AAK33713.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 68 AA; 7790 MW; C2E2D3DA9412A754 CRC64;  
 SO  
 Query Match 100.0%; Score 21; DB 5; Length 68;  
 Best Local Similarity 100.0%; Pred. No. 6.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 RGDA 4  
 ||||  
 Db 41 RGDA 44

RESULT 15  
 Q8UJK6 PRELIMINARY; PRT; 68 AA.  
 ID Q8UJK6  
 AC Q8UJK6;  
 DT 01-JUN-2002 (TEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)  
 DE Hypothetical protein Atu5470.  
 GN ATU5470 OR AGR PAT 693.  
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
 OC Plasmid AT.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Rhizobium.  
 OX NCBI\_TaxID=176299;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21608550; PubMed=11743193;  
 RA Wood D.W., Sebda J.C., Kaul R., Monks D.E., Kitajima J.P.,  
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,  
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bove D. Sr.,  
 RA Chapman P., Glendinning J., Deatherage G., Gillet W., Grant C.,  
 RA Kuyavin T., Levy R., Li M.-J., Kocelland E., Palmeri A.,  
 RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,  
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,  
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,  
 RA Nestor E.W.;  
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens  
 RT C58.";  
 RL Science 294:2317-2323 (2001).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21608551; PubMed=11743194;  
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
 RA Querillo B., Goldman B.S., Cao Y., Askenazi M., Hailing C., Mullin L.,  
 RA Houmlet K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,  
 RA Houtan C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,  
 RA Flanagan C., Crowell C., Guron J., Lomo C., Sear C., Strub G.,  
 RA Cielo C., Slater S.;  
 RT "Genome sequence of the plant pathogen and biotechnology agent  
 RT Agrobacterium tumefaciens C58.";  
 RL Science 294:2323-2328 (2001).  
 DR EMBL; AE008968; AAL46157.1; -.  
 DR EMBL; AE007916; AAK90845.1; -.  
 KW Hypothetical protein; Plasmid; Complete proteome.  
 SQ SEQUENCE 68 AA; 8005 MW; 5CABE406D75E93A8 CRC64;  
 SO  
 Query Match 100.0%; Score 21; DB 16; Length 68;  
 Best Local Similarity 100.0%; Pred. No. 6.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 RGDA 4  
 ||||  
 Db 36 RGDA 39  
 Search completed: February 11, 2004, 14:56:02  
 Job time : 9.83671 secs

OK protein - protein search, using SW model

Run on: February 11, 2004, 14:35:52 ; Search time 25.9335 Seconds

(without alignments)  
73.441 Million cell updates/sec

Title: US-10-050-611-2

Perfect score: 69

Sequence: 1 DACEGDSGPFV 12

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726373 residues

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*

2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*

3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*

4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*

5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*

6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*

7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*

8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*

9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*

10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*

11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*

12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*

13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*

14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*

15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*

16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*

17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*

18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*

19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*

20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*

21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*

22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*

23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

ALIGNMENTS

Result No.	Score	Query Match	Length	DB ID	Description
1	69	100.0	12	23	AAW50857
2	69	100.0	23	20	AAW83414
3	69	100.0	23	21	AAW12893
4	69	100.0	23	22	AAW70363
5	69	100.0	23	23	AAW22563
6	69	100.0	23	23	AAW20159
7	69	100.0	23	23	AAW78376
8	69	100.0	23	23	AAW50858
9	69	100.0	23	24	AAW72755
10	69	100.0	23	24	AAW72757
11	69	100.0	23	24	AAW72758
12	69	100.0	33	24	AAW72758
13	69	100.0	111	20	AAW99113
14	69	100.0	116	20	AAW9115
15	69	100.0	259	18	AAW11545
16	69	100.0	259	24	AAW60563
17	69	100.0	259	24	AAW60565
18	69	100.0	295	16	AAW74775
19	69	100.0	295	16	AAW74776
20	69	100.0	295	16	AAW74777
21	69	100.0	295	16	AAW74778
22	69	100.0	295	16	AAW74779
23	69	100.0	295	16	AAW74780
24	69	100.0	295	16	AAW76033
25	69	100.0	295	16	AAW76034
26	69	100.0	295	16	AAW76035
27	69	100.0	295	16	AAW76036
28	69	100.0	295	16	AAW76037
29	69	100.0	295	16	AAW76038
30	69	100.0	295	16	AAW76039
31	69	100.0	295	16	AAW76040
32	69	100.0	295	18	AAW22892
33	69	100.0	295	21	AAW08613
34	69	100.0	295	24	AAW60562
35	69	100.0	295	24	AAW60564
36	69	100.0	308	20	AAW99107
37	69	100.0	308	20	AAW99109
38	69	100.0	376	14	AAW41797
39	69	100.0	376	20	AAW42789
40	69	100.0	376	23	AAW10703
41	69	100.0	579	14	AAW35763
42	69	100.0	579	18	AAW11546
43	69	100.0	579	18	AAW11544
44	69	100.0	579	20	AAW9108
45	69	100.0	582	20	AAW99106

RESULT 1  
 ID AAM50857  
 AA AAM50857 standard; Peptide; 12 AA.  
 XX  
 AC AAM50857;  
 XX  
 DT 01-MAY-2002 (first entry)  
 XX  
 DE Serine esterase conserved sequence used in cardiac tissue repair.  
 XX  
 KW Serine esterase; thrombin; revascularisation; vascular occlusion;  
 KW tissue repair; valvulopathy; vasculopathy; cardiac; angiogenesis;  
 KW restenosis; therapy; enzyme; human.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200204008-A2.  
 XX  
 PD 17-JAN-2002.  
 XX  
 PF 12-JUL-2001; 2001WO-US21944.  
 XX  
 PR 12-JUL-2000; 2000US-217583P.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 XX  
 PI Carney DH;  
 XX  
 DR WPI; 2002-179665/23.  
 XX  
 PT Promoting cardiac tissue repair, stimulating revascularisation,  
 PT stimulating vascular endothelial cell proliferation, and inhibiting  
 PT vascular occlusion by using angiogenic thrombin derivative peptide -  
 XX  
 PS Claim 3; Page 19; 24pp; English.  
 XX  
 CC The present peptide comprises a thrombin-derived serine esterase  
 CC conserved sequence that is used in a claimed method for promoting  
 CC cardiac tissue repair. The method involves administering an  
 CC angiogenic thrombin-derived peptide, especially a thrombin receptor  
 CC binding domain comprising the 4-amino acid peptide given in  
 CC AAM50856 together with the serine esterase conserved sequence,  
 CC or preferably the peptide given in AAM50858, which includes both  
 CC these peptide sequences. The thrombin-derived peptide is  
 CC administered during or following cardiac surgery by injection  
 CC into cardiac tissue, and may be formulated as a sustained release  
 CC formulation. It is used in claimed methods of stimulating  
 CC revascularisation, stimulating vascular endothelial cell  
 CC proliferation, inhibiting vascular occlusion, and inhibiting  
 CC restenosis following balloon angioplasty, in which case the  
 CC peptide may be coated onto the catheter.  
 XX  
 SQ Sequence 12 AA;  
 Query Match 100.0%; Score 69; DB 23; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.0029;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGGGFV 12  
 |||||  
 Db 1 DACEGDSGGGFV 12  
 RESULT 2  
 ID AAM83414  
 AA AAM83414 standard; peptide; 23 AA.  
 XX  
 AC AAM83414;  
 XX  
 DT 26-FEB-1999 (first entry)  
 XX  
 DE Cell growth/adhesion promoting peptide #1.  
 XX  
 KW Cell growth; adhesion; promotion; medical treatment; injury;  
 KW biotissue; bone reinforcement; nerve regeneration; BMP resin.  
 XX  
 OS Synthetic.  
 XX  
 PN JP10316581-A.  
 XX  
 PD 02-DEC-1998.  
 XX  
 PF 15-MAY-1997; 97JP-0140885.  
 XX  
 PR 15-MAY-1997; 97JP-0140885.  
 XX  
 PA (KURS ) KURARAY CO LTD.  
 XX  
 DR WPI; 1999-076400/07.  
 XX  
 PT Material for medical treatment comprises new peptide - used for  
 PT covering injuries, promoting adhesion of bio-tissues, bone  
 PT reinforcing and nerve regeneration  
 XX  
 PS Claim 1; Page 12; 14pp; Japanese.  
 XX  
 CC The present invention describes a material for medical treatment which  
 CC comprises one or more peptides of the formula XADBCGTLPRQY, or their  
 CC salts, immobilised on a substrate; where X = H, CH3CO or CH3COLys;  
 CC A = Ser or Thr; D = Ile, Val or Leu; E = Lys or Arg; G = Ile, Val or  
 CC Leu; J = Gly or Ala; L = Ile, Val or Leu; M = Gly or Ala; Q = Gly, Ala  
 CC or Gly-Lys-Lys-Gly; Y = OH or NH2. Also described is an agent for cell  
 CC growth promotion and/or cell adhesion promotion containing the above  
 CC peptide or its salt as the active component. The peptide and its salt  
 CC can be used for covering injuries, promoting adhesion of biotissues,  
 CC bone reinforcing and nerve regeneration. The present sequence represents  
 CC a specifically claimed peptide of the present invention.  
 XX  
 SQ Sequence 23 AA;  
 Query Match 100.0%; Score 69; DB 20; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 0.0051;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGGPFV 12  
DB 12 DACEGDSGGPFV 23

RESULT 3  
AAB12893  
ID AAB12893 standard; peptide; 23 AA.

AC AAB12893;

DT 02-NOV-2000 (first entry)

DE Nerve tissue regenerative peptide SEQ ID #8.

KW Nerve regeneration; nerve cell proliferation; axon extension; treatment; central nervous system disorder; peripheral nervous system disorder;

KW spinal disorder; head injury; cerebrovascular disorder.

OS Synthetic.

PN JP2000143531-A.

PD 23-MAY-2000.

PF 11-AUG-1999; 99JP-0227108.

PR 09-SEP-1998; 98JP-0270498.

PA (KURSA) KURARAY CO LTD.

PA (NISHI) NISHIMURA Y.

PA (SUZU) SUZUKI Y.

PA (TANI) TANIHARA M.

DR WPI; 2000-415772/36.

PT New nerve regeneration material -

PS Claim 2; Page 5; 17pp; Japanese.  
CC This invention relates to a new nerve regenerative material which  
CC contains a peptide immobilised to a base which consists of a  
CC polysaccharide gel such as alginate acid. Sequences AAB12886-B12899  
CC represent examples of the peptides used in the nerve regeneration  
CC material. The peptide containing material causes nerve cell  
CC proliferation and also causes axonal extension. The material can be used  
CC for the treatment of central or peripheral nervous system disorders,  
CC spinal disorders, head injury or cerebrovascular disorders.

SQ Sequence 23 AA;

Query Match 100.0%; Score 69; DB 21; Length 23;  
Best Local Similarity 100.0%; Pred. No. 0.0051;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGGPFV 12  
DB 12 DACEGDSGGPFV 23

DB 12 DACEGDSGGPFV 23  
RESULT 4  
AAB70363  
ID AAB70363 standard; peptide; 23 AA.

AC AAB70363;

DT 02-MAY-2001 (first entry)

DE Human thrombin receptor binding domain peptide SPQ ID NO:8.

KW Neutrophil cell chemotactic; wound healing; inflammation; vulnery;

KW antiinflammatory.

OS Homo sapiens.

PN US6184342-B1.

PD 06-FEB-2001.

PF 28-OCT-1994; 94US-0330594.

PR 28-OCT-1994; 94US-0330594.

PA (CHRY-) CHRYSALIS BIOTECHNOLOGY INC.

PA Carney DH, Ramakrishnan S;

DR WPI; 2001-202003/20.  
PT New synthetic neutrophil cell chemotactic peptides, useful for  
PT generating antibodies for modulating neutrophil chemotaxis in immune  
PT response and wound healing -

PS Example 2; Column 6; 15pp; English.  
CC The present invention describes a synthetic peptide (I) which is a  
CC neutrophil cell chemotactic agent. (I) has vulnerary and  
CC antiinflammatory activities. (I) is useful as a potent neutrophil cell  
CC chemotactic agent and for generating antibodies against the peptides,  
CC which are useful for modulating neutrophil recruitment to a wound site  
CC for enhancing or inhibiting inflammation and early effects of wound  
CC healing. Neutrophil response to (I) is specific, since monocytes and  
CC fibroblasts do not show any expression of the receptor to which (I)  
CC binds. The present sequence represents a human thrombin receptor binding  
CC domain peptide which is used in an example from the present invention.

SQ Sequence 23 AA;

Query Match 100.0%; Score 69; DB 22; Length 23;  
Best Local Similarity 100.0%; Pred. No. 0.0051;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGGPFV 12  
DB 12 DACEGDSGGPFV 23



Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGGPFV 12  
 |||||  
 Db 12 DACEGDSGGPFV 23

RESULT 7  
 AAU78376  
 ID AAU78376 standard; Peptide; 23 AA.  
 XX  
 AC AAU78376;  
 DT 18-JUN-2002 (first entry)  
 DE Thrombin peptide derivative TP508.  
 KW Thrombin; osteopathic; receptor; agonist; bone growth stimulation;  
 KW osteoinduction; farm animal; companion animal; laboratory animal;  
 KW bone graft; segmental bone gap; bone void; non-union fracture.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Misc-difference 3 /label= Unknown  
 XX  
 XX WO200205836-A2.  
 XX 24-JAN-2002.  
 XX 18-JUL-2001; 2001WO-US22641.  
 XX 19-JUL-2000; 2000US-219300P.  
 XX (TEXA ) UNIV TEXAS SYSTEM.  
 PI Carney DH, Crowther RS, Simmons DJ, Yang J, Redin WR;  
 DR WPI; 2002-303796/34.  
 XX  
 PT Stimulating bone growth at a site in a subject in need of  
 PT osteoinduction, such as a site of bone graft, segmental bone gap, bone  
 PT void or non-union structure, by administering agonist of activated  
 PT thrombin receptor -  
 XX  
 PS Claim 11; Page 22; 27pp; English.  
 CC The invention describes a method of stimulating bone growth at a site  
 CC in a subject in need of osteoinduction. The method involves administering  
 CC an agonist to stimulate bone growth at a site in a subject (e.g. a farm  
 CC animal, companion animal or laboratory animal), in need of  
 CC osteoinduction, such as the site in need of a bone graft in a subject, a  
 CC segmental bone gap, a bone void or a non-union fracture. This sequence  
 CC represents a thrombin peptide derivative obtained from a serine esterase  
 CC that can stimulate or activate the non-proteolytically activated thrombin  
 CC receptor.

XX SQ Sequence 23 AA;  
 Query Match 100.0%; Score 69; DB 23; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 0.0051;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGGPFV 12  
 |||||  
 Db 12 DACEGDSGGPFV 23

RESULT 8  
 AAU50856  
 ID AAU50856 standard; Peptide; 23 AA.  
 XX  
 AC AAU50856;  
 DT 01-MAY-2002 (first entry)  
 DE Thrombin-derived peptide used to promote cardiac tissue repair.  
 KW Thrombin; revascularisation; vascular occlusion; tissue repair;  
 KW vulnerable; vasotropic; cardiac; angiogenesis; restenosis;  
 KW therapy; human.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Peptide 10..13 /note= "thrombin receptor binding domain"  
 FT Peptide 12..23 /note= "serine esterase conserved sequence"  
 XX  
 XX WO200204008-A2.  
 XX 17-JAN-2002.  
 XX 12-JUL-2001; 2001WO-US21944.  
 XX 12-JUL-2000; 2000US-217583P.  
 XX (TEXA ) UNIV TEXAS SYSTEM.  
 PI Carney DH;  
 DR WPI; 2002-179665/23.  
 CC Promoting cardiac tissue repair, stimulating revascularisation,  
 CC stimulating vascular endothelial cell proliferation, and inhibiting  
 CC vascular occlusion by using angiogenic thrombin derivative peptide -  
 PS Claim 4; Page 19; 24pp; English.  
 CC The present peptide comprises a thrombin-derived peptide, TP508,  
 CC that includes a thrombin receptor binding domain sequence (see also  
 CC AAU50856) and a serine esterase conserved sequence (see also

CC AAM50837). The peptide is used in a claimed method for promoting  
 CC cardiac tissue repair. It is administered during or following  
 CC cardiac surgery by injection into cardiac tissue, and may be  
 CC formulated as a sustained release formulation. The thrombin  
 CC derivative peptide is also used in claimed methods of stimulating  
 CC revascularization, stimulating vascular endothelial cell  
 CC proliferation, inhibiting vascular occlusion, and inhibiting  
 CC restenosis following balloon angioplasty, in which case it may be  
 CC coated onto the catheter.  
 CC  
 SQ Sequence 23 AA;  
 QY Query Match 100.0%; Score 69; DB 23; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 0.0051;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 DB 1 DACEGDSGGPFV 12  
 12 DACEGDSGGPFV 23  
 RESULT 9  
 ABP72755  
 ID ABP72755 standard; Peptide; 23 AA.  
 AC  
 XX ABP72755;  
 DT 11-JUN-2003 (first entry)  
 DE Antilucer peptide derived from human thrombin.  
 XX  
 KW Antilucer; human; thrombin.  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal H or R3-C(O), where R3 is H or  
 FT a Cl-C6 alkyl group"  
 FT Misc-difference 3 /note= "given as Try in the specification"  
 FT Modified-site 23  
 FT /note= "C-terminal OH or NR4R5, where R4 and R5 are  
 FT independently H, a Cl-C6 alkyl group or,  
 FT taken together with the N atom to which they  
 FT are bonded, a non-aromatic heterocyclic  
 FT group"  
 FT Modified-site 1..23  
 FT /note= "Q, 1, 2 or 3 amino acids at positions 1-9  
 FT and 14-23 differ from the given sequence  
 FT e.g. are conservative substitutions of the  
 FT amino acid at the corresponding position of  
 FT this sequence"  
 XX  
 XX  
 PN WO2003013569-A2.  
 XX

PD 20-FEB-2003.  
 XX  
 PF 16-JAN-2002; 2002WO-US01151.  
 XX  
 PR 27-JUL-2001; 2001US-308198P.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 XX  
 PI Carney DH;  
 PI  
 DR WPI; 2003-289898/28.  
 XX  
 PT Promoting healing of chronic dermal skin ulcer such as diabetic ulcer,  
 PT on a subject, by contacting the skin ulcer with an agonist of  
 PT non-proteolytically activated thrombin receptor -  
 PS Claim 1; Page 14; 19pp; English.  
 XX  
 CC The present sequence is that of a human thrombin-derived peptide  
 CC based on prothrombin amino acid residues 508-530. The peptide acts  
 CC as an agonist of the non-proteolytically activated thrombin  
 CC receptor and has antilucer activity. A claimed method of promoting  
 CC healing of a chronic dermal skin ulcer on a subject comprises  
 CC contacting the ulcer with an effective amount of this peptide, or an  
 CC N-terminal truncated fragment of it having at least 14 amino acids,  
 CC or a C-terminal truncated fragment of it having at least 18 amino  
 CC acids. Preferably, the peptide has -H at the N-terminus and -NH2 or  
 CC -OH at the C-terminus. An example is peptide TP508 (see ABP72757),  
 CC which was shown in an example from the invention to accelerate  
 CC the healing of chronic diabetic ulcers and to increase the  
 CC percentage of ulcer closure. The thrombin-derived peptides of the  
 CC invention can be used to treat a chronic dermal skin ulcer,  
 CC especially a diabetic ulcer, decubitus ulcer, venous stasis ulcer  
 CC or an arterial ulcer on a human, a companion animal, farm animal or  
 CC laboratory animal. They are inexpensive to produce and cause few,  
 CC if any, side effects.  
 XX  
 SQ Sequence 23 AA;  
 QY Query Match 100.0%; Score 69; DB 24; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 0.0051;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 DB 1 DACEGDSGGPFV 12  
 12 DACEGDSGGPFV 23  
 RESULT 10  
 ABP72757  
 ID ABP72757 standard; Peptide; 23 AA.  
 AC  
 XX ABP72757;  
 DT 11-JUN-2003 (first entry)  
 DE Antilucer peptide TP508 derived from human thrombin.  
 XX  
 XX

XX Antilucer; human; thrombin.  
 KW  
 XX Homo sapiens.  
 OS  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 3 /note= "given as Try in the specification"  
 FT Modified-site 23 /note= "C-terminal amide"  
 FT  
 XX  
 PN WO2003013569-A2.  
 XX  
 PD 20-FEB-2003.  
 XX  
 PF 16-JAN-2002; 2002WO-US01151.  
 XX  
 PR 27-JUL-2001; 2001US-308198P.  
 XX  
 XX (TEXA ) UNIV TEXAS SYSTEM.  
 PA  
 PI Carney DH;  
 XX  
 PI WPI; 2003-289898/28.  
 DR  
 XX  
 PT Promoting healing of chronic dermal skin ulcer such as diabetic ulcer,  
 PT on a subject, by contacting the skin ulcer with an agonist of  
 PT non-proteolytically activated thrombin receptor -  
 PT  
 XX  
 PS Claim 15; Page 16; 19pp; English.  
 XX  
 CC The present sequence is that of a preferred human thrombin-derived  
 CC peptide of the invention that is based on prothrombin amino acid  
 CC residues 508-530. It is denoted TP508. The peptide acts as an  
 CC agonist of the non-proteolytically activated thrombin receptor and  
 CC has antilucer activity. In an example from the invention, TP508  
 CC is shown to accelerate the healing of chronic diabetic ulcers and  
 CC to increase the percentage of ulcer closure. The antilucer  
 CC peptides of the invention can be used to treat a chronic dermal  
 CC skin ulcer, especially a diabetic ulcer, decubitus ulcer, venous  
 CC stasis ulcer or an arterial ulcer on a human, a companion animal,  
 CC farm animal or laboratory animal. The peptides are inexpensive to  
 CC produce and cause few, if any, side effects.  
 CC  
 XX  
 SQ Sequence 23 AA;  
 Query Match 100.0%; Score 69; DB 24; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 0.0051;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ABP72760  
 ID ABP72760 standard; Peptide; 23 AA.  
 XX  
 AC ABP72760;  
 XX  
 XX 11-JUN-2003 (first entry)  
 DE Human thrombin peptide fragment.  
 XX  
 KW Antilucer; human; thrombin.  
 XX  
 OS Homo sapiens.  
 OS  
 FH Key Location/Qualifiers  
 FT Misc-difference 3 /note= "given as Try in the specification"  
 FT  
 XX  
 PN WO2003013569-A2.  
 XX  
 PD 20-FEB-2003.  
 XX  
 PF 16-JAN-2002; 2002WO-US01151.  
 XX  
 PR 27-JUL-2001; 2001US-308198P.  
 XX  
 XX (TEXA ) UNIV TEXAS SYSTEM.  
 PA  
 PI Carney DH;  
 XX  
 PI WPI; 2003-289898/28.  
 DR  
 XX  
 PT Promoting healing of chronic dermal skin ulcer such as diabetic ulcer,  
 PT on a subject, by contacting the skin ulcer with an agonist of  
 PT non-proteolytically activated thrombin receptor -  
 PT  
 XX  
 PS Disclosure; Page 3; 19pp; English.  
 XX  
 CC The present sequence is that of a human thrombin-derived peptide  
 CC comprising prothrombin amino acid residues 508-530. The invention  
 CC provides peptides based on this sequence (see ABP72755-59) that act  
 CC as agonists of the non-proteolytically activated thrombin receptor  
 CC and which have antilucer activity. One of these thrombin-derived  
 CC peptides (see ABP72756) was shown to accelerate the healing of  
 CC chronic diabetic ulcers and to increase the percentage of ulcer  
 CC closure. The peptides of the invention can be used to treat a  
 CC chronic dermal skin ulcer, especially a diabetic ulcer, decubitus  
 CC ulcer, venous stasis ulcer or an arterial ulcer on a human, a  
 CC companion animal, farm animal or laboratory animal. They are  
 CC inexpensive to produce and cause few, if any, side effects.  
 CC  
 XX  
 SQ Sequence 23 AA;  
 Query Match 100.0%; Score 69; DB 24; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 0.0051;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



DR WPI; 1999-070237/06.

XX Exosite assay for agents that inhibit catalytic cleavage of

PT prothrombin - at sites away from the active site of prothrombinase,

PT also new inhibitors, potentially useful as anticoagulants

XX

PS Disclosure; Page 42-43; 61pp; English.

XX An exosite assay has been developed for inhibition of the catalytic

CC cleavage of prothrombin (Pth) to thrombin (Th) by prothrombinase (I), at

CC a site remote from the catalytic site of (I) comprises: (a) preparing a

CC solution containing 0.05-20  $\mu$ M substrate (S), that includes a protease

CC cleavage site and exosite-binding determinant; 0.05-200 nM factor Va;

CC 30-500 micro M phospholipids (PL); test inhibitor (A) in buffer of pH

CC 7-9, containing 1-10 mM calcium ions but no calcium-chelating agent;

CC (b) initiating catalytic cleavage of (S) by adding an aliquot of factor

CC Xa (to final concentration 0.05-200 nM) so that there is an excess of Va

CC over Xa, forming a S/(I) complex; (c) withdrawing aliquots of the

CC reaction mixture, quenching them; and (d) assaying for concentration of

CC Th. Alternatively, in the initial solution S is replaced by the same

CC concentration of Xa (less than the amount of Va), and reaction is started

CC by adding S. Also described in the present invention are inhibitors (A') are

CC having IC50 less than 1  $\mu$ M identified by this assay. (A') are

CC potentially useful as a new class of anticoagulants for treatment of

CC cardiovascular disease, stroke and hematological disorders. The method

CC is based on the finding that exosite interactions are critical for

CC substrate specificity in catalytic formation of Th. The present sequence

CC represents bovine zeta 2 prothrombin 2.

CC

XX

SQ Sequence 111 AA;

QY Query Match 100.0%; Score 69; DB 20; Length 111;

Best Local Similarity 100.0%; Pred. No. 0.021;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 DACEGDSGGPFV 12

51 DACEGDSGGPFV 62

RESULT 14

AAW9115

ID AAW9115 standard; protein; 116 AA.

XX

AC AAW9115;

XX

DT 14-MAY-1999 (first entry)

XX

DE Human zeta 2 prothrombin 2.

XX

XX Prothrombin; exosite assay; anticoagulant; blood clot; thrombin;

KW cardiovascular disease; stroke; hematological disorder.

XX

OS Homo sapiens.

XX

PN W09653130-A1.

XX

PD 10-DEC-1998.

XX

PF 28-MAY-1998; 98WO-US10840.

XX

XX 08-APR-1998; 98US-0081030.

PR 06-JUN-1997; 97US-0048864.

XX

PA (VEM-) UNIV EMORY.

XX

PI Krishnaswamy S;

XX

XX WPI; 1999-070237/06.

DR

XX Exosite assay for agents that inhibit catalytic cleavage of

PT prothrombin - at sites away from the active site of prothrombinase,

PT also new inhibitors, potentially useful as anticoagulants

XX

PS Disclosure; Page 44-45; 61pp; English.

XX An exosite assay has been developed for inhibition of the catalytic

CC cleavage of prothrombin (Pth) to thrombin (Th) by prothrombinase (I), at

CC a site remote from the catalytic site of (I) comprises: (a) preparing a

CC solution containing 0.05-20  $\mu$ M substrate (S), that includes a protease

CC cleavage site and exosite-binding determinant; 0.05-200 nM factor Va;

CC 30-500 micro M phospholipids (PL); test inhibitor (A) in buffer of pH

CC 7-9, containing 1-10 mM calcium ions but no calcium-chelating agent;

CC (b) initiating catalytic cleavage of (S) by adding an aliquot of factor

CC Xa (to final concentration 0.05-200 nM) so that there is an excess of Va

CC over Xa, forming a S/(I) complex; (c) withdrawing aliquots of the

CC reaction mixture, quenching them; and (d) assaying for concentration of

CC Th. Alternatively, in the initial solution S is replaced by the same

CC concentration of Xa (less than the amount of Va), and reaction is started

CC by adding S. Also described in the present invention are inhibitors (A') are

CC having IC50 less than 1  $\mu$ M identified by this assay. (A') are

CC potentially useful as a new class of anticoagulants for treatment of

CC cardiovascular disease, stroke and hematological disorders. The method

CC is based on the finding that exosite interactions are critical for

CC substrate specificity in catalytic formation of Th. The present sequence

CC represents human zeta 2 prothrombin 2.

CC

XX

SQ Sequence 116 AA;

QY Query Match 100.0%; Score 69; DB 20; Length 116;

Best Local Similarity 100.0%; Pred. No. 0.021;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 DACEGDSGGPFV 12

56 DACEGDSGGPFV 67

RESULT 15

AAW11545

ID AAW11545 standard; Protein; 259 AA.

XX

AC AAW11545;

XX

DT 01-OCT-1997 (first entry)  
 XX Human thrombin Asn99 mutant.  
 XX  
 KM Prothrombin; mutant; mutase; platelet aggregation; blood clotting;  
 KM coagulation; reduce; decrease; hirudin; heparin; anti-thrombin III;  
 KM antagonist; D99N.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Protein 1..259  
 FT /label= thrombin\_Asn99  
 FT Misc-difference 99  
 FT /note= "Wild-type Asp residue has been replaced by  
 FT Asn"  
 XX  
 PN WO9641865-A2.  
 XX  
 PD 27-DEC-1996.  
 XX  
 PE 12-JUN-1996; 96WO-AT00105.  
 XX  
 PR 13-JUN-1995; 95AT-0001006.  
 XX  
 PA (IMMO ) IMMO AG.  
 XX  
 PI Eibl J, Falkner F, Fischer B, Mitterer A, Schlokat U;  
 XX WPI; 1997-065455/06.  
 DR  
 XX  
 FT Prothrombin mutants with reduced clotting activity - useful as  
 FT antagonists of thrombin inhibitors or for anticoagulant therapy  
 XX  
 PS Example 3; Page -; 73pp; German.  
 XX  
 CC Prothrombin mutants having one or more changes in amino acid sequence  
 CC compared with the natural protein and having 0-10% (preferably 0-0.25%)  
 CC of the activity of the natural protein are claimed, provided that the  
 CC changes in amino acid sequence do not affect the capacity of the  
 CC mutants to bind to specific ligands and receptors. The mutants have  
 CC greatly reduced clotting activity and are useful as antagonists of  
 CC thrombin inhibitors such as hirudin, heparin and anti-thrombin III.  
 CC The mutations may also result in changes to the in vivo half-life  
 CC of prothrombin. The half-life may be reduced to less than 10 minutes  
 CC or the mutant prothrombin may have an extended half-life of more than  
 CC 1 hour, making it useful as an anticoagulant and to inhibit side-  
 CC effects of anti-coagulant treatment. They are converted to inactive  
 CC thrombin and are able to compete with native, active thrombin for  
 CC binding to receptors. The present sequence represents the thrombin  
 CC mutant which is derived by trypsin cleavage of a specifically  
 CC claimed human prothrombin mutant in which Asp at position 419 is  
 CC changed to Asn. The thrombin Asn99 mutant was found to have only  
 CC 0.24% of the activity of wild-type thrombin on a chromogenic  
 CC substrate.  
 CC (Note: This sequence does not appear in the specification and has

CC been produced by modifying the wild-type sequence of human  
 CC prothrombin which appears in figure 1).  
 XX  
 SQ Sequence 259 AA;  
 QY 1 DACEGDSGGPFY 12  
 Db 199 DACEGDSGGPFY 210  
 Query Match 100.0%; Score 69; DB 18; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 0.044;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Search completed: February 11, 2004, 14:53:24  
 Job time : 25.9355 secs

OK protein - protein search, using SW model

Run on: February 11, 2004, 14:49:07 ; Search time 8.12903 Seconds  
(without alignments)  
141.963 Million cell updates/sec

Title: US-10-050-611-2  
Perfect score: 69  
Sequence: 1 DACEGDSGGPFV 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616662 residues  
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69	100.0	234	2 F42696	thrombin (EC 3.4.21.5)
2	69	100.0	235	2 D42696	thrombin (EC 3.4.21.5)
3	69	100.0	235	2 F42696	thrombin (EC 3.4.21.5)
4	69	100.0	236	2 C42696	thrombin (EC 3.4.21.5)
5	69	100.0	236	2 I42696	thrombin (EC 3.4.21.5)
6	69	100.0	239	2 G42696	thrombin (EC 3.4.21.5)
7	69	100.0	617	2 S10511	thrombin (EC 3.4.21.5)
8	69	100.0	618	2 A35827	thrombin (EC 3.4.21.5)
9	69	100.0	622	1 T8HJ	thrombin (EC 3.4.21.5)
10	69	100.0	625	1 T8BO	thrombin (EC 3.4.21.5)
11	66	95.7	417	1 S00845	hepsin (EC 3.4.21.5)
12	66	95.7	1524	2 T30357	polyprotein - Afri
13	63	91.3	235	2 H42696	thrombin (EC 3.4.21.5)

14	63	91.3	456	1 KXBO	protein C (activat
15	63	91.3	461	1 KXHU	protein C (activat
16	60	87.0	254	2 S69465	trypsin-like prote
17	60	87.0	256	1 TRFE	trypsin-like prote
18	60	87.0	264	2 S3794	trypsin-like prote
19	60	87.0	267	2 S40006	trypsin (EC 3.4.21
20	60	87.0	271	2 S41308	serine proteinase
21	60	87.0	274	2 S35339	trypsin (EC 3.4.21
22	60	87.0	275	2 S40007	trypsin (EC 3.4.21
23	60	87.0	275	2 S40005	trypsin (EC 3.4.21
24	60	87.0	277	2 S39340	trypsin (EC 3.4.21
25	60	87.0	285	2 T35195	probable serine pr
26	60	87.0	394	2 J50600	t-plasminogen acti
27	60	87.0	431	2 J50599	t-plasminogen acti
28	60	87.0	461	1 S18994	protein C (activat
29	60	87.0	461	1 JX0210	protein C (activat
30	60	87.0	477	1 A34369	t-plasminogen acti
31	60	87.0	477	2 J50597	t-plasminogen acti
32	60	87.0	477	2 J50598	t-plasminogen acti
33	60	87.0	559	1 A35029	t-plasminogen acti
34	60	87.0	559	1 A28941	t-plasminogen acti
35	60	87.0	562	1 UKHUT	t-plasminogen acti
36	60	87.0	593	2 S45281	coagulation factor
37	60	87.0	603	2 S28941	coagulation factor
38	60	87.0	615	1 KFRU12	coagulation factor
39	59	85.5	161	2 I62744	coagulation factor
40	59	85.5	161	2 I48158	coagulation factor
41	59	85.5	265	2 I15451	hypothetical prote
42	59	85.5	275	2 I46712	factor IX - rabbit
43	59	85.5	282	2 I84621	coagulation factor
44	59	85.5	434	1 A35005	u-plasminogen acti
45	59	85.5	459	2 J00419	coagulation factor

#### ALIGNMENTS

RESULT 1  
F42696  
thrombin (EC 3.4.21.5) B chain - Cynops pyrogastor (fire-bellied newt)  
(fragment)  
C/Species: Cynops pyrogastor (fire-bellied newt)  
C/Date: 19-Mar-1997 #sequence\_revision 19-Dec-1997 #text\_change 17-Mar-1999  
C/Accession: F42696  
R/Banfield, D.K.; MacGillivray, R.T.A.  
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992  
A/Title: Partial characterization of verterbrate prothrombin cDNAs: amplification  
and sequence analysis of the B chain of thrombin from nine different species.  
A/Reference number: A42696; WUID:92212913; PMID:1557363  
A/Note: sequence not  
A/Accession: F42696  
A/Status: preliminary; nucleic acid sequence not shown; not compared with  
conceptual translation  
A/Molecule type: RNA  
A/Residues: 1-234 <BAN>  
A/Cross-references: GB:M81395  
C/Superfamily: thrombin; Gla domain homology; kingle homology; trypsin homology

C/Keywords: hydrolase; serine proteinase

Query Match 100.0%; Score 69; DB 2; Length 234;

Best Local Similarity 100.0%; Pred. No. 0.00051;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGGPFV 12

|||||

Db 174 DACEGDSGGPFV 185

## RESULT 2

D42696

thrombin (EC 3.4.21.5) B chain - chicken (fragment)

C/Species: Gallus gallus (chicken)

C/Date: 26-May-1994 #sequence\_revision 26-May-1994 #text\_change 17-Mar-1999

C/Accession: D42696

R/Banfield, D.K.; MacGillivray, R.T.A.

Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992

A/Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and sequence analysis of the B chain of thrombin from nine different species.

A/Reference number: A42696; PMID:1597383

A/Accession: D42696

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-235 <BAN>

A/Cross-references: GB:M81391

C/Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology

C/Keywords: hydrolase; serine proteinase

F/1-226/Domain: trypsin homology (fragment) <TRY>

Query Match 100.0%; Score 69; DB 2; Length 235;

Best Local Similarity 100.0%; Pred. No. 0.00052;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGGPFV 12

|||||

Db 175 DACEGDSGGPFV 186

## RESULT 3

E42696

thrombin (EC 3.4.21.5) B chain - tokay (fragment)

C/Species: Gekko gekko (tokay)

C/Date: 26-May-1994 #sequence\_revision 26-May-1994 #text\_change 17-Mar-1999

C/Accession: E42696

R/Banfield, D.K.; MacGillivray, R.T.A.

Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992

A/Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and sequence analysis of the B chain of thrombin from nine different species.

A/Reference number: A42696; PMID:1597383

A/Accession: E42696

A/Status: preliminary; nucleic acid sequence not shown; not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 1-235 <BAN>

A/Cross-references: GB:M81392

C/Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology

C/Keywords: hydrolase; serine proteinase

F/1-226/Domain: trypsin homology (fragment) <TRY>

Query Match 100.0%; Score 69; DB 2; Length 235;

Best Local Similarity 100.0%; Pred. No. 0.00052;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGGPFV 12

|||||

Db 175 DACEGDSGGPFV 186

## RESULT 4

C42696

thrombin (EC 3.4.21.5) B chain - rabbit (fragment)

C/Species: Oryctolagus cuniculus (domestic rabbit)

C/Date: 26-May-1994 #sequence\_revision 26-May-1994 #text\_change 17-Mar-1999

C/Accession: C42696

R/Banfield, D.K.; MacGillivray, R.T.A.

Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992

A/Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and sequence analysis of the B chain of thrombin from nine different species.

A/Reference number: A42696; PMID:1597383

A/Accession: C42696

A/Status: preliminary; nucleic acid sequence not shown; not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 1-236 <BAN>

A/Cross-references: GB:M81396

C/Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology

C/Keywords: hydrolase; serine proteinase

F/1-227/Domain: trypsin homology (fragment) <TRY>

Query Match 100.0%; Score 69; DB 2; Length 236;

Best Local Similarity 100.0%; Pred. No. 0.00052;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGGPFV 12

|||||

Db 176 DACEGDSGGPFV 187

## RESULT 5

I42696

thrombin (EC 3.4.21.5) B chain - Pacific hagfish (fragment)

C/Species: Eptatretus stouti (Pacific hagfish)

C/Date: 26-May-1994 #sequence\_revision 26-May-1994 #text\_change 17-Mar-1999

C/Accession: I42696

R/Banfield, D.K.; MacGillivray, R.T.A.

Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992

A/Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and sequence analysis of the B chain of thrombin from nine different species.

A/Reference number: A42696; PMID:1597383

A/Accession: I42696

A/Status: preliminary; not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 1-236 <BAN>  
 A/Cross-references: GB:M81393  
 A/Note: nucleotide translation not given  
 C/Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology  
 C/Keywords: hydrolase; serine proteinase  
 F/1-26/Domain: trypsin homology (fragment) <TRY>

Query Match 100.0%; Score 69; DB 2; Length 236;  
 Best Local Similarity 100.0%; Pred. No. 0.00052;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGSPFV 12  
 |||||||||||  
 Db 175 DACEGDSGSPFV 186

RESULT 6

G42696  
 thrombin (EC 3.4.21.5) B chain - rainbow trout (fragment).  
 C/Species: Oncorhynchus mykiss (rainbow trout)  
 C/Date: 26-May-1994 #sequence\_revision 26-May-1994 #text\_change 22-Jun-1999  
 C/Accession: G42696  
 R/Banfield, D.K.; Macgillivray, R.T.A.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992  
 A/Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and sequence analysis of the B chain of thrombin from nine different species.  
 A/Reference number: A42696; MUID:92212913; PMID:1557383  
 A/Accession: G42696  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-239 <BAN>  
 A/Cross-references: GB:M81398; NID:q213486; PIND:AAA49433.1; PID:q213487  
 C/Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology  
 C/Keywords: hydrolase; serine proteinase  
 F/1-26/Domain: trypsin homology (fragment) <TRY>

Query Match 100.0%; Score 69; DB 2; Length 239;  
 Best Local Similarity 100.0%; Pred. No. 0.00052;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGSPFV 12  
 |||||||||||  
 Db 175 DACEGDSGSPFV 186

RESULT 7

S10511  
 thrombin (EC 3.4.21.5) Precursor - rat  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 03-May-2002  
 C/Accession: S10511; A60576; B42696  
 R/Dhanich, M.; Monard, D.  
 Nucleic Acids Res. 18, 4251, 1990  
 A/Title: cDNA sequence of rat prothrombin.  
 A/Reference number: S10511; MUID:90332426; PMID:2377469  
 A/Accession: S10511  
 A/Molecule type: mRNA

A/Residues: 1-617 <DIR>  
 A/Cross-references: EMBL:X52835; NID:q56969; PIND:CAA37017.1; PID:q56970  
 R/Henriksson, K.P.; Jazin, E.E.; Greenwood, J.A.; Dickerman, H.W.  
 Endocrinology 126, 167-175, 1990  
 A/Title: Prothrombin levels are increased in the estrogen-treated immature rat uterus.  
 A/Reference number: A60576; MUID:90091942; PMID:2293980  
 A/Accession: A60576  
 A/Molecule type: protein  
 A/Residues: 44-58 <HEN>  
 A/Note: the authors purified the proenzyme from the estrogen-stimulated maturing rat uterus and demonstrated it to be prothrombin  
 R/Banfield, D.K.; Macgillivray, R.T.A.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992  
 A/Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and sequence analysis of the B chain of thrombin from nine different species.  
 A/Reference number: A42696; MUID:92212913; PMID:1557383  
 A/Accession: B42696  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 383-617, 'E' <BAN>  
 A/Cross-references: GB:M81397  
 C/Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology  
 C/Keywords: blood coagulation; calcium binding; carboxylutamic acid; glycoprotein; hydrolase; kringle; serine proteinase  
 F/1-24/Domain: signal sequence #status predicted <SIG>  
 F/25-43/Domain: propeptide #status predicted <PRO>  
 F/48-88/Domain: Gla domain homology <GLA>  
 F/44-617/Product: prothrombin #status experimental <PMAT>  
 F/109-187/Domain: kringle homology <KR1>  
 F/215-292/Domain: kringle homology <KR2>  
 F/360-609/Domain: trypsin homology <TRY>  
 F/50, 51, 58, 63, 64, 69, 70, 73, 76/Modified site: gamma-carboxylutamic acid (Glu)  
 #status predicted  
 F/61-66, 91-104, 109-187, 130-170, 158-182, 215-292, 236-276, 264-287, 332-478, 387-403, 532-546, 560-590/Disulfide bonds: #status predicted  
 F/402, 458, 564/Active site: His, Asp, Ser #status predicted

Query Match 100.0%; Score 69; DB 2; Length 617;  
 Best Local Similarity 100.0%; Pred. No. 0.0013;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGSPFV 12  
 |||||||||||  
 Db 558 DACEGDSGSPFV 569

RESULT 8

A35827  
 thrombin (EC 3.4.21.5) precursor - mouse  
 C/Species: Mus musculus (house mouse)  
 C/Date: 14-Dec-1990 #sequence\_revision 14-Dec-1990 #text\_change 03-May-2002  
 C/Accession: A35827; A42696; S12081  
 R/Degen, S.J.F.; Schaefer, L.A.; Jamison, C.S.; Grant, S.G.; Fitzgerald, J.J.; Pal, J.A.; Chapman, V.M.; Elliott, R.W.  
 DNA Cell Biol. 9, 487-498, 1990

A>Title: Characterization of the cDNA coding for mouse prothrombin and localization of the gene on mouse chromosome 2.  
A/Reference number: A35827; MUID:91025551; PMID:2222810  
A/Accession: A35827  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-618 <DEG>  
A/Cross-references: GB:M52308; NID:953813; PIDN:CAA3548.1; PID:953814  
A/Experimental source: strain C57BL/6  
A/Note: the data were obtained from females resulting from the cross of M. domesticus and M. spretus  
R:Bantfield, D.K.; Macgillivray, R.T.A.  
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992  
A>Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and sequence analysis of the B chain of thrombin from nine different species.  
A/Reference number: A42696; MUID:92212913; PMID:1557383  
A/Accession: A42696  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 384-618, 'E' <BAN>  
A/Cross-references: GB:M61394  
C/Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology  
C/Keywords: blood coagulation; calcium binding; carboxylglutamic acid; glycoprotein; hydrolase; kringle; serine protease  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:28-48/Domain: propeptide #status predicted <PRO>  
F:44-618/Product: prothrombin B #status predicted <MAT>  
F:109-187/Domain: kringle homology <KR1>  
F:215-293/Domain: kringle homology <KR2>  
F:361-610/Domain: trypsin homology <TRY>  
F:50,51,58,60,63,64,69,70,73,76/Modified site: gamma-carboxylglutamic acid (Glu)  
#status predicted  
F:61-66,91-104,109-187,130-170,158-182,215-293,236-276,264-288,333-479,388-404,533-547,561-591/Disulfide bonds: #status predicted  
F:403,459,565/Active site: His, Asp, Ser #status predicted

Query Match 100.0%; Score 69; DB 2; Length 618;  
Best Local Similarity 100.0%; Pred. No. 0.0013;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGPFV 12  
|||||  
DB 559 DACEGDSGPFV 570

RESULT 9  
1BHU  
thrombin (EC 3.4.21.3) precursor [validated] - human  
N/Alternate names: coagulation factor II  
N/Contents: prothrombin  
C/Species: Homo sapiens (man)  
C/Date: 30-Nov-1980 #sequence revision 22-Jul-1994 #text change 08-Dec-2000  
C/Accession: A29351; A00914; B00914; A37549; A37550; I5152  
R/Degen, S.J.F.; Davie, E.W.  
Biochemistry 26, 6165-6177, 1987  
A>Title: Nucleotide sequence of the gene for human prothrombin.

A/Reference number: A29351; MUID:88077877; PMID:2825773  
A/Accession: A29351  
A/Molecule type: DNA  
A/Residues: 1-622 <DEG>  
A/Cross-references: GB:M17262; GB:M33691; NID:9558069; PIDN:AA63054.1; PID:933961  
R/Degen, S.J.F.; Macgillivray, R.T.A.; Davie, E.W.  
Biochemistry 22, 2087-2097, 1983  
A>Title: Characterization of the complementary deoxyribonucleic acid and gene coding for human prothrombin.  
A/Reference number: A00914; MUID:83231469; PMID:6305407  
A/Accession: A00914  
A/Molecule type: mRNA  
A/Residues: 8-163, 'N', '165-622 <DE2>  
A/Cross-references: GB:V00595; GB:J00307; NID:937128; PIDN:CAA23842.1; PID:91335344  
A/Accession: B00914  
A/Molecule type: DNA  
A/Residues: 158-311 <DE3>  
R/Waltz, D.A.; Hewett-Ewart, D.; Seegers, W.H.  
Proc. Natl. Acad. Sci. U.S.A. 74, 1969-1972, 1977  
A/Reference number: A37549; MUID:77193964; PMID:266717  
A/Accession: A37549  
A/Molecule type: protein  
A/Residues: 44-118, 'N', '120', '15', '122-163', '1', '165-175', 'A', '177-182', 'T', '184-193', 'M', '196-308', 'E', '309-314 <MAL>  
R/Bukowski, R.J.; Ellison, J.; Downing, M.R.; Mann, K.G.  
J. Biol. Chem. 252, 4942-4957, 1977  
A>Title: Primary structure of human prothrombin 2 and alpha-thrombin.  
A/Reference number: A37550; MUID:77207112; PMID:873923  
A/Accession: A37550  
A/Molecule type: protein  
A/Residues: 315-354, 'N', '336-348', 'N', '350-366', 'N', '370-397', 'N', '399-413', 'N', '415-464', 'N', '486-493', 'G', '495-503', 'V', '505-508', 'S', '510', 'V', '512-513', 'D', '519-528', 'AL', '531', 'O', '533-622 <BUT>  
R/Rablier, M.J.; Blaehill, A.; Furie, B.; Furie, B.C.  
J. Biol. Chem. 261, 13210-13215, 1986  
A/Reference number: A37551; MUID:87068332; PMID:3759958  
A/Contents: annotation; activation cleavages  
R/Macgillivray, R.T.; Irvine, D.M.; Quinto, E.R.; Stone, J.C.  
Ann. N.Y. Acad. Sci. 485, 73-79, 1986  
A>Title: Recombinant genetic approaches to functional mapping of thrombin.  
A/Reference number: I51952; MUID:87182874; PMID:3471191  
A/Accession: I51952  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-2, 'R', '5-100 <RES>  
A/Cross-references: GB:M33031; NID:9190723; PIDN:AAA60220.1; PID:9190724  
C/Comment: Thrombin, which cleaves bonds after Arg and Lys, converts fibrinogen to fibrin and activates factors V, VIII, XIII, and, in complex with thrombomodulin, protein C.  
C/Comment: Prothrombin is activated on the surface of a phospholipid membrane that binds the amino end of prothrombin and factors Va and Xa in calcium-dependent interactions. The activation peptide(s) can be removed either by factor Xa or thrombin; the cleavage into light and heavy chains is by factor Xa. It is not known whether one or two smaller activation peptides, with additional cleavage after 314-Arg, are released in natural blood clotting.

C/Comment: The cleavage after Arg-198, observed in vitro, does not occur in plasma.

C/Comment: The gamma-carboxyglutamyl residues bind calcium ions, result from the carboxylation of glutamyl residues by microsomal vitamin K-dependent carboxylase, and are necessary for calcium-dependent interaction with the negatively charged phospholipid membrane surface.

C/Comment: The prothrombin precursor is synthesized in the liver.

C/Genetics:

A/Genes: GDB:F2

A/Cross-references: GDB:19894; OMIM:176930

A/Map position: 11p11-11q12

A/Introns: 27/1; 80/3; 89/1; 106/1; 141/2; 187/1; 292/1; 335/1; 377/2; 433/2; 491/2; 552/1; 575/3

C/Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology

C/Keywords: acute phase; blood coagulation; calcium binding; carboxyglutamic acid; duplication; glycoprotein; hydrolase; kringle; liver; plasma; serine proteinase

F/1-24/Domain: signal sequence #status predicted <SIG>

F/25-43/Domain: propeptide #status predicted <PRO>

F/28-87/Domain: Gla domain homology <GLA>

F/44-622/Product: prothrombin #status experimental <MAT>

F/108-186/Domain: activation peptide #status experimental <APT>

F/213-291/Domain: kringle homology <KR1>

F/328-363/Product: thrombin light chain #status experimental <LC>

F/364-622/Product: thrombin heavy chain #status experimental <HC>

F/49,50,57,59,62,63,68,69,72,75/Modified site: gamma-carboxyglutamic acid (Glu)

#status experimental

F/60-65,90-103,108-186,189-157,181,213-291,234-274,262-286/Disulfide bonds: #status predicted

F/121,143/Binding site: carboxylate (Asn) (covalent) #status predicted

F/336-482,536-550,564-594/Disulfide bonds: #status predicted

F/391-407/Disulfide bonds: #status experimental

F/406,462/Active site: His, Asp #status predicted

F/416/Binding site: carboxylate (Asn) (covalent) #status experimental

F/568/Active site: Ser #status experimental

Query Match 100.0%; Score 69; DB 1; Length 622;

Best Local Similarity 100.0%; Pred. No. 0.0013;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGGPFV 12

Do 562 DACEGDSGGPFV 573

RESULT 10

TBBO

Thrombin (EC 3.4.21.5) precursor - bovine

C/Species: Bos primigenius taurus (cattle)

C/Date: 24-Apr-1984 #sequence revision 14-Jul-1994 #text\_change 18-Jun-1999

C/Accession: S02537; A00915; A37552; I46045; S67518

R/Title: Thrombin, which cleaves bonds after Arg and Lys, converts fibrinogen to fibrin and activates factors V, VII, VIII, XIII, and, in complex with thrombomodulin, protein C.

J. Mol. Biol. 200, 31-45, 1988

A/Title: Structure and evolution of the bovine prothrombin gene.

A/Reference number: S02537; MUID:88245190; PMID:3379642

A/Accession: S02537

A/Status: not compared with conceptual translation

A/Molecule type: DNA

A/Residues: 1-625 <IRM>

R/MacGillivray, R.T.A.; Davie, E.W.

Biochemistry 23, 1626-1634, 1984

A/Title: Characterization of bovine prothrombin mRNA and its translation product.

A/Reference number: A00915; MUID:84203525; PMID:6326805

A/Accession: A00915

A/Molecule type: mRNA

A/Residues: 1-230, 'H', 232-625 <MXC>

A/Note: 600-Asn was also found

R/Magnusson, S.; Sottcup-Jensen, L.; Petersen, T.E.; Claess, H.

In Boehringer Symposium on Prothrombin and Related Coagulation Factors, Henker, H.C., and Velkamp, J.J., eds., pp.25-46, Leiden Univ. Press, Leiden, 1975

A/Reference number: A37552

A/Accession: A37552

A/Molecule type: protein

A/Residues: 44-287, 'N', 289-352, 'E', 354, 'Q', 356-548, 'ND', 551-599, 'N', 601-625 <MXG>

A/Note: the evidence for 231-Ser is strong

A/Note: disulfide bonds and carboxylate binding sites were determined

R/Park, C.H.; Tulinsky, A.

Biochemistry 25, 3977-3982, 1986

A/Title: Three-dimensional structure of the kringle sequence: structure of prothrombin fragment 1

A/Reference number: A37553; MUID:86296631; PMID:3741841

A/Contents: annotation; residues 44-317, X-ray crystallography, 2.8 angstroms

R/Ritlin, D.M.; Ahern, K.G.; Pearson, G.D.; MacGillivray, R.T.A.

Biochemistry 24, 6854-6861, 1985

A/Title: Characterization of the bovine prothrombin gene.

A/Reference number: A37554; MUID:86077733; PMID:3000440

A/Contents: annotation; gene structure

R/MacGillivray, R.T.; Degen, S.J.; Chandra, T.; Woo, S.L.; Davie, E.W.

Proc. Natl. Acad. Sci. U.S.A. 77, 5153-5157, 1980

A/Title: Cloning and analysis of a cDNA coding for bovine prothrombin.

A/Reference number: I46045; MUID:81054926; PMID:6254059

A/Accession: I46045

A/Status: preliminary; translated from GE/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 466-599, 'N', 601-625 <MA2>

A/Cross-references: EMBL:V00135; NID:9772; PIND:CAA23451.1; PID:9808945

R/Pejler, G.; Karlstrom, A.R.; Berg, L.

Eur. J. Biochem. 227, 102-107, 1995

A/Title: Identification of the proteolytic thrombin fragments formed after cleavage with rat mast cell protease 1.

A/Reference number: S67518; MUID:95154277; PMID:7851376

A/Accession: S67518

A/Status: preliminary

A/Molecule type: protein

A/Residues: 318-325,333-338, 'X', 340,367-374,481-484, 'X', 486-488,515-522 <PEJ>

C/Comment: Thrombin, which cleaves bonds after Arg and Lys, converts fibrinogen to fibrin and activates factors V, VII, VIII, XIII, and, in complex with thrombomodulin, protein C.

A/Comment: Prothrombin is activated on the surface of a phospholipid membrane that binds the amino end of prothrombin and factors Va and Xa in calcium-

dependent interactions; factor Xa removes the activation peptide and cleaves the remaining part into light and heavy chains. The activation process starts slowly because factor V itself has to be activated by the initial, small amounts of thrombin.

C/Comment: Thrombin can cleave the amino-terminal activation peptide 1 from prothrombin, prior to its activation by factor Xa.

C/Comment: The gamma-carboxyglutamate residues bind calcium ions, result from the carboxylation of glutamate residues by microsomal vitamin K-dependent carboxylase, and are necessary for calcium-dependent interaction with the negatively charged phospholipid membrane surface.

C/Comment: The prothrombin precursor is synthesized in the liver.

C/Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology

C/Keywords: blood coagulation; calcium binding; carboxyglutamic acid;

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-43/Domain: propeptide #status predicted <PRO>

F:28-88/Domain: Gla domain homology <GLA>

F:44-625/Product: prothrombin #status experimental <PRT>

F:44-199/Domain: activation peptide 1 #status experimental <R1>

F:109-187/Domain: kringle homology <KR1>

F:200-317/Domain: activation peptide 2 #status experimental <FR2>

F:214-292/Domain: kringle homology <KR2>

F:318-366/Product: thrombin light chain #status experimental <LCH>

F:367-625/Product: thrombin heavy chain #status experimental <HC>

F:367-616/Domain: trypsin homology <TRY>

F:750-51,56,60,63,64,69,70,73,76/Modified site: gamma-carboxyglutamic acid (Glu)

#status experimental

F:61-66,91-104,109-187,130-170,158-182,214-292,235-275,263-287,339-485,394-

410,535-553,567-597/Dissulfide bonds: #status experimental

F:120,144,419/Binding site: carboxylate (Asn) (covalent) #status experimental

F:409,465,571/Active site: His, Asp, Ser #status experimental

Query Match 100.0%; Score 69; DB 1; Length 625;

Best Local Similarity 100.0%; Pred. No. 0.0013;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGGPFV 12

DB 565 DACEGDSGGPFV 576

RESULT 11

S00845

hepsin (EC 3.4.21.-) - human

C/Species: Homo sapiens (man)

C/Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 18-Jun-1999

C/Accession: S00845

Rileykus, S.P.; Loeb, K.R.; Hagen, F.S.; Kurachi, K.; Davie, E.W.

Biochemistry 27, 1067-1074, 1988

A/Title: A novel trypsin-like serine protease (hepsin) with a putative transmembrane domain expressed by human liver and hepatoma cells.

A/Reference number: S00845; PMID:88209431; PMID:2835076

A/Accession: S00845

A/Molecule type: mRNA

A/Residues: 1-417 <LEV>

A/Cross-references: EMBL:X07732; NID:932063; PID:CAA30558.1; PID:932064

C/Genetics:

A/Gene: GDB:HPN; TMRSS1; hepsin

A/Cross-references: GDB:135685; OMIM:142440

A/Map position: 19q11-19q13.2

C/Superfamily: hepsin; trypsin homology

C/Keywords: hydrolase; liver; serine protease; transmembrane protein

F:23-45/Domain: transmembrane #status predicted <TM>

F:163-400/Domain: trypsin homology <TRY>

F:188-204,291-359,322-338,349-381/Dissulfide bonds: #status predicted

F:203,257,353/Active site: His, Asp, Ser #status predicted

Query Match 95.7%; Score 66; DB 1; Length 417;

Best Local Similarity 91.7%; Pred. No. 0.0028;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGGPFV 12

DB 347 DACEGDSGGPFV 358

RESULT 12

T30337

polyprotein - African clawed frog

C/Species: Xenopus laevis (African clawed frog)

C/Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 03-Feb-2003

C/Accession: T30337

RiYang, J.C.; Lindsay, L.L.; Hedrick, J.L.

submitted to the EMBL Data Library, March 1998

A/Description: cDNA cloning of ovocytinase, a chymotrypsin-like protease released from Xenopus laevis eggs at fertilization.

A/Reference number: 220829

A/Accession: T30337

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-1524 <YAN>

A/Cross-references: EMBL:U81290; NID:92981640; PID:92981641; PID:AA024717.1

C/Superfamily: tyrosin related polypeptide; trypsin homology

Query Match 95.7%; Score 66; DB 2; Length 1524;

Best Local Similarity 91.7%; Pred. No. 0.0093;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGGPFV 12

DB 241 DACEGDSGGPFV 252

Search completed: February 11, 2004, 14:56:56

Job time: 8.12903 secs

OM protein - protein search, using sw model

Run on: February 11, 2004, 14:36:52 ; Search time 5.03226 Seconds  
(without alignments)  
112.141 Million cell updates/sec

Title: US-10-050-611-2  
Perfect score: 69  
Sequence: 1 DACEGDSGGPFLV 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues  
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution..

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	69	100.0	617	1	THRB_RAT
2	69	100.0	618	1	THRB_MOUSE
3	69	100.0	622	1	THRB_HUMAN
4	69	100.0	625	1	THRB_BOVIN
5	66	95.7	417	1	HEPS_HUMAN
6	66	95.7	436	1	HEPS_MOUSE
7	63	91.3	157	1	PRTC_CANFA
8	63	91.3	157	1	PRTC_CAPII
9	63	91.3	157	1	PRTC_FELCA
10	63	91.3	157	1	PRTC_HORSE
11	63	91.3	161	1	PRTC_MACMU
12	63	91.3	456	1	PRTC_BOVIN
13	63	91.3	459	1	PRTC_PIG
14	63	91.3	461	1	PRTC_HUMAN
15	60	87.0	248	1	KIKC_HUMAN
16	60	87.0	253	1	TRVB_DROER
17	60	87.0	253	1	TRVD_DROER

18	60	87.0	253	1	TRVD_DROME	P42276 drosophila
19	60	87.0	253	1	TRVG_DROME	P42277 drosophila
20	60	87.0	254	1	TRVP_SASBU	P13388 sarcophaga
21	60	87.0	256	1	HYPB_HYPLI	P35368 hypoderma 1
22	60	87.0	256	1	TRVA_DROER	P54624 drosophila
23	60	87.0	256	1	TRVA_DROME	P04814 drosophila
24	60	87.0	256	1	TRVE_DROME	P54627 drosophila
25	60	87.0	256	1	TRVE_DROME	P35005 drosophila
26	60	87.0	258	1	TRVD_DROER	P54629 drosophila
27	60	87.0	262	1	TRVU_DROME	P42279 drosophila
28	60	87.0	264	1	VDE_BOMMO	Q07943 bombyx mori
29	60	87.0	267	1	TRV7_ANGGA	P35041 anopheles g
30	60	87.0	274	1	TRV1_ANGGA	P35035 anopheles g
31	60	87.0	275	1	TRV3_ANGGA	P35037 anopheles g
32	60	87.0	275	1	TRV4_ANGGA	P35038 anopheles g
33	60	87.0	277	1	KLMD_HUMAN	Q9UKC3 homo sapien
34	60	87.0	277	1	TRV2_ANGGA	P35036 anopheles g
35	60	87.0	281	1	TRVZ_DROER	P54630 drosophila
36	60	87.0	394	1	URTG_DESRO	P49150 desmodus ro
37	60	87.0	418	1	HATT_HUMAN	O60235 homo sapien
38	60	87.0	422	1	DESL_HUMAN	Q9U152 homo sapien
39	60	87.0	431	1	URTB_DESRO	P98121 desmodus ro
40	60	87.0	455	1	TMS5_MOUSE	Q9EFC4 mus musculus
41	60	87.0	457	1	TMS5_HUMAN	Q9U350 homo sapien
42	60	87.0	458	1	PRTC_RABIT	Q28661 oryctolagus
43	60	87.0	461	1	PRTC_MOUSE	P33587 mus musculus
44	60	87.0	461	1	PRTC_RAT	P31394 rattus norv
45	60	87.0	477	1	URTI_DESRO	P98119 desmodus ro

ALIGNMENTS

RESULT 1	ID	THRB_RAT	STANDARD	PRT	617 AA.
AC	P18292;				
DT	01-NOV-1990 (Rel. 16, Created)				
DT	01-NOV-1990 (Rel. 16, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Prothrombin precursor (EC 3.4.21.5).				
GN					
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=10116;				
[1]					
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Sprague-Dawley; TISSUE=Liver;				
RX	MEDLINE=90332426; PubMed=2377469;				
RA	Dhanraj M., Monard D.;				
RT	"cDNA sequence of rat prothrombin."				
RL	Nucleic Acids Res. 18:4251-4251(1990).				
RN	[2]				
RP	SEQUENCE OF 383-617 FROM N.A.				
RC	TISSUE=Liver;				
RX	MEDLINE=92212913; PubMed=1557383;				

RA Banfield D.K., Macgillivray R.T.:  
 RT "Partial characterization of vertebrate prothrombin cDNAs:  
 RT amplification and sequence analysis of the B chain of thrombin from  
 RT nine different species."  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).  
 CC -1- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS  
 CC FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII,  
 CC AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.  
 CC -1- CATALYTIC ACTIVITY: preferential cleavage: Arg-1-Gly; activates  
 CC fibrinogen to fibrin and releases fibrinopeptide A and B.  
 CC -1- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,  
 CC RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOXAL  
 CC ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES  
 CC ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY  
 CC CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION  
 CC OF PROTHROMBIN TO THROMBIN.  
 CC -1- MISCELLANEOUS: PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A  
 CC PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN &  
 CC FACTORS VA & XA IN CA-DEPENDENT INTERACTIONS; FACTOR XA REMOVES  
 CC THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT &  
 CC HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR  
 CC V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF  
 CC THROMBIN.  
 CC -1- MISCELLANEOUS: THROMBIN CAN ITSELF CLEAVE THE AMINO TERMINAL  
 CC FRAGMENT (FRAGMENT 1) OF THE PROTHROMBIN, PRIOR TO ITS ACTIVATION  
 CC BY FACTOR XA.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -1- SIMILARITY: Contains 2 kringle domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 DR EMBL: X52835; CAA37017.1; -.  
 DR EMBL: M81397; AAA42240.1; -.  
 DR PIR: S10511; S10511.  
 DR HSSP: P00734; IUVS.  
 DR MEROPS: S01.217; -.  
 DR InterPro: IPR001314; Chymotrypsin.  
 DR InterPro: IPR002383; GLA\_blood.  
 DR InterPro: IPR000001; Kringle.  
 DR InterPro: IPR003666; Prothrombin.  
 DR InterPro: IPR001254; Ser protease\_Try.  
 DR InterPro: IPR000294; Vitr\_dep\_GLA.  
 DR Pfam: PF00594; gla; 1.  
 DR Pfam: PF00051; Kringle\_2.  
 DR Pfam: PF00089; trypsin\_1.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR PRINTS: PR00001; GLABLOOD.  
 DR PRINTS: PR00018; KRINGLE.  
 DR PRINTS: PR01505; PROTHROMBIN.  
 DR ProDom: PD000395; Kringle\_2.  
 DR SMART: SM00069; GLA\_1.

DR SMART: SM00130; KR; 2.  
 DR SMART: SM00020; try\_spec; 1.  
 DR PROSITE: PS00011; GLU\_CARBOXYLATION; 1.  
 DR PROSITE: PS00021; KRINGLE\_1; 2.  
 DR PROSITE: PS50070; KRINGLE\_2; 2.  
 DR PROSITE: PS50240; TRYPsin\_DOM; 1.  
 DR PROSITE: PS00134; TRYPsin\_HIS; 1.  
 DR PROSITE: PS00135; TRYPsin\_SER; 1.  
 DR Vitamin K; Zymogen; Gamma-carboxyglutamic acid; Acute phase; Liver;  
 DR Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;  
 DR Hydrilase; Serine protease; Kringle; Signal.  
 FT SIGNAL 1 24  
 FT PROPER 25 43  
 FT CHAIN 44 617  
 FT PEPTIDE 44 200  
 FT PEPTIDE 201 323  
 FT CHAIN 324 359  
 FT CHAIN 360 617  
 FT DOKAIN 109 187  
 FT DOKAIN 215 292  
 FT DOKAIN 360 617  
 FT SITE 200 201  
 FT SITE 323 324  
 FT SITE 359 360  
 FT ACT SITE 402 402  
 FT ACT SITE 458 458  
 FT ACT SITE 564 564  
 FT MOD RES 50 50  
 FT MOD RES 51 51  
 FT MOD RES 58 58  
 FT MOD RES 60 60  
 FT MOD RES 63 63  
 FT MOD RES 64 64  
 FT MOD RES 69 69  
 FT MOD RES 70 70  
 FT MOD RES 73 73  
 FT MOD RES 76 76  
 FT CARBOHYD 120 120  
 FT CARBOHYD 144 144  
 FT CARBOHYD 412 412  
 FT CARBOHYD 552 552  
 FT DISULFID 61 66  
 FT DISULFID 91 104  
 FT DISULFID 109 187  
 FT DISULFID 130 170  
 FT DISULFID 158 182  
 FT DISULFID 215 292  
 FT DISULFID 236 276  
 FT DISULFID 264 287  
 FT DISULFID 332 478  
 FT DISULFID 387 403  
 FT DISULFID 532 546  
 FT DISULFID 560 590  
 SQ SEQUENCE 617 AA; 70411 MW; AD27D1B71445B3.D CRC64;  
 Query Match 100.0%; Score 69; DB 1; Length 617;  
 Best Local Similarity 100.0%; Pred. No. 0.00031;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGGPFV 12  
 558 DACEGDSGGPFV 569

Do

RESULT 2

THROMBIN\_MOUSE STANDARD; PRT; 618 AA.

AC P19221;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE Prothrombin precursor (EC 3.4.21.5).  
 OS F2 OR CF2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Liver;  
 RX MEDLINE=91025551; PubMed=2222810;  
 RA Fritzer Degen S.O.; Schaffer L.A.; Jamison C.S.; Grant S.G.;  
 RA Fitzgibbon J.J.; Pal J.-A.; Chapman V.K.; Elliott R.W.;  
 RA "Characterization of the cDNA coding for mouse prothrombin and  
 RT localization of the gene on mouse chromosome 2.";  
 RL DNA Cell Biol. 9:487-498(1990).  
 RN [2]  
 RP SEQUENCE OF 384-618 FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=92212913; PubMed=1557383;  
 RA Banfield D.K.; Macgillivray R.T.;  
 RT "Partial characterization of vertebrate prothrombin cDNAs:  
 RT amplification and sequence analysis of the B chain of thrombin from  
 RT nine different species.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).  
 CC -1- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS  
 CC FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII,  
 CC AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.  
 CC -1- CATALYTIC ACTIVITY: preferential cleavage: Arg-|-Gly; activates  
 CC fibrinogen to fibrin and releases fibrinopeptide A and B.  
 CC -1- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,  
 CC RESULT FROM THE CARBOXYGLUTAMYL RESIDUES BY A MICROSOMAL  
 CC ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES  
 CC ARE NECESSARY FOR THE Ca-DEPENDENT INTERACTION WITH A NEGATIVELY  
 CC CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION  
 CC OF PROTHROMBIN TO THROMBIN.  
 CC -1- MISCELLANEOUS: PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A  
 CC PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN &  
 CC FACTORS VA & XA IN Ca-DEPENDENT INTERACTIONS. FACTOR XA REMOVES  
 CC THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT &  
 CC HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR  
 CC V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF  
 CC THROMBIN.  
 CC -1- MISCELLANEOUS: THROMBIN CAN ITSELF CLEAVE THE AMINO TERMINAL

CC FRAGMENT (FRAGMENT 1) OF THE PROTHROMBIN, PRIOR TO ITS ACTIVATION  
 CC BY FACTOR XA.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -1- SIMILARITY: Contains 2 kringle domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 CC EMBL; X52308; AAA36548.1; -.  
 CC EMBL; M81394; AAA40435.1; -.  
 CC PIR; A35827; A35827.  
 CC HSP; P00734; 1B7X.  
 CC MEPS; S01.217; -.  
 CC MED; MG; 88380; F2.  
 CC InterPro; IPR001314; Chymotrypsin.  
 CC InterPro; IPR002383; GLA\_Blood.  
 CC InterPro; IPR000001; Kringle.  
 CC InterPro; IPR003966; Prothrombin.  
 CC InterPro; IPR001254; Ser. protease\_Try.  
 CC InterPro; IPR000294; Vltk\_deg\_GLA.  
 CC Pfam; PF00594; GLA; 1.  
 CC Pfam; PF00053; Kringle; 2.  
 CC Pfam; PF00089; trypsin; 1.  
 CC PRINTS; PR00722; CHYMOTRYPSIN.  
 CC PRINTS; PR00001; GLABLOOD.  
 CC PRINTS; PR00018; KRINGLE.  
 CC PRINTS; PR01505; PROTHROMBIN.  
 CC ProDom; PD000395; Kringle; 2.  
 CC SMART; SM00069; GLA; 1.  
 CC SMART; SM00130; KR; 2.  
 CC SMART; SM00020; TRY\_Spe; 1.  
 CC PROSITE; PS00011; GLU CARBOXYLATION; 1.  
 CC PROSITE; PS00021; KRINGLE\_1; 2.  
 CC PROSITE; PS00070; KRINGLE\_2; 2.  
 CC PROSITE; PS0240; TRYPsin\_DOM; 1.  
 CC PROSITE; PS00134; TRYPsin\_HIS; 1.  
 CC PROSITE; PS00135; TRYPsin\_SER; 1.  
 CC Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;  
 CC Vitamin K; Zymogen; gamma-carboxyglutamic acid; Acute phase; Liver;  
 CC Hydrolyase; Serine protease; Kringle; Signal.  
 CC SIGNAL  
 CC FT 1 24  
 CC FT PROPEP 25 43  
 CC FT CHAIN 44 618  
 CC FT PEPTIDE 44 200  
 CC FT PEPTIDE 201 324  
 CC FT CHAIN 325 360  
 CC FT CHAIN 361 618  
 CC FT DOMAIN 109 187  
 CC FT DOMAIN 215 292  
 CC FT DOMAIN 361 618  
 CC FT SITE 200 201  
 CC FT SITE 324 325  
 CC PROTHROMBIN.  
 CC ACTIVATION PEPTIDE (FRAGMENT 1).  
 CC THROMBIN LIGHT CHAIN (A).  
 CC THROMBIN HEAVY CHAIN (B).  
 CC KRINGLE 1.  
 CC KRINGLE 2.  
 CC SERINE PROTEASE.  
 CC CLEAVAGE (BY THROMBIN).  
 CC CLEAVAGE (BY FACTOR XA).

FT	SITE	360	361	CLEAVAGE (BY FACTOR XA).
FT	ACT SITE	403	403	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT SITE	459	459	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT SITE	565	565	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	MOD_RES	50	50	GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES	51	51	GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES	58	58	GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES	60	60	GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES	63	63	GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES	64	64	GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES	69	69	GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES	70	70	GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES	73	73	GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES	76	76	GAMMA-CARBOXYGLUTAMIC ACID.
FT	DISULFID	61	66	BY SIMILARITY.
FT	DISULFID	91	104	BY SIMILARITY.
FT	DISULFID	109	187	BY SIMILARITY.
FT	DISULFID	130	170	BY SIMILARITY.
FT	DISULFID	158	182	BY SIMILARITY.
FT	DISULFID	215	293	BY SIMILARITY.
FT	DISULFID	236	276	BY SIMILARITY.
FT	DISULFID	264	288	BY SIMILARITY.
FT	DISULFID	333	479	INTERCHAIN (BY SIMILARITY).
FT	DISULFID	388	404	BY SIMILARITY.
FT	DISULFID	533	547	BY SIMILARITY.
FT	DISULFID	561	591	BY SIMILARITY.
FT	CARBOHYD	122	122	N-LINKED (GLCNAC. .).
FT	CARBOHYD	144	144	N-LINKED (GLCNAC. .).
FT	CARBOHYD	413	413	N-LINKED (GLCNAC. .).
FT	CARBOHYD	553	553	N-LINKED (GLCNAC. .).
SC	SEQUENCE	618 AA;	70268 MW;	B89F719A9FD601E0 CRC64;

Query Match 100.0%; Score 69; DB 1; Length 618;  
 Best Local Similarity 100.0%; Pred. No. 0.00031;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DACEGDSGPFV 12  
 |||||  
 Db 559 DACEGDSGPFV 570

RESULT 3  
 THRB HUMAN  
 ID THRB\_HUMAN STANDARD; PRT; 622 AA.  
 AC P00734;  
 DT 01-JUN-1986 (Rel. 01, Created)  
 DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DE Prothrombin precursor (EC 3.4.21.5) (Coagulation factor II).  
 GN F2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88077877; PubMed=2825773;

RA Degen S.J.F., Davie E.W.;  
 RT "Nucleotide sequence of the gene for human prothrombin.";  
 RL Biochemistry 26:6165-6177(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND VARIANT MET-165.  
 RA Rieder M.J., Arnel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,  
 RA Ozuna M., Poel C.L., Toth E.J., Yi Q., Nickerson D.A.;  
 RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 8-622 FROM N.A.  
 RA MEDLINE=83231469; PubMed=6305407;  
 RA Degen S.J.F., McGilivray R.T.A., Davie E.W.;  
 RT "Characterization of the complementary deoxyribonucleic acid and gene  
 RT coding for human prothrombin.";  
 RL Biochemistry 22:2087-2097(1983).  
 RN [4]  
 RP SEQUENCE OF 44-314.  
 RA MEDLINE=77193964; PubMed=266717;  
 RA Walz D.A., Hewett-Emslett D., Seegers W.H.;  
 RT "Amino acid sequence of human prothrombin fragments 1 and 2.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 74:1969-1972(1977).  
 RN [5]  
 RP SEQUENCE OF 315-622.  
 RA MEDLINE=77207112; PubMed=873923;  
 RA Burkowski R.J., Elion J., Downing M.R., Mann K.G.;  
 RT "Primary structure of human prothrombin 2 and alpha-thrombin.";  
 RL J. Biol. Chem. 252:4942-4957(1977).  
 RN [6]  
 RP PROCESSING.  
 RA MEDLINE=87008532; PubMed=3759958;  
 RA Rabiet M.J., Blashill A., Furie B., Furie B.C.;  
 RT "Prothrombin fragment 1 X 2 X 3, a major product of prothrombin  
 RT activation in human plasma.";  
 RL J. Biol. Chem. 261:13210-13215(1986).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
 RA MEDLINE=90059942; PubMed=2583108;  
 RA Bode W., Mayr I., Baumann U., Huber R., Stone S.R., Hofsteenge J.;  
 RT "The refined 1.9 A crystal structure of human alpha-thrombin:  
 RT interaction with D-Phe-Pro-Arg chloromethylketone and significance of  
 RT the Tyr-Pro-Tyr insertion segment.";  
 RL EMBO J. 8:3467-3475(1989).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RA MEDLINE=90327074; PubMed=2374926;  
 RA Rydel T.J., Ravichandran K.G., Tulinsky A., Bode W., Huber R.,  
 RA Roitsch C., Fenton J.W. II;  
 RT "The structure of a complex of recombinant hirudin and human alpha-  
 RT thrombin.";  
 RL Science 249:277-280(1990).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
 RA MEDLINE=94350942; PubMed=8071320;  
 RA Rydel T.J., Yin M., Padmanabhan K.P., Blankenship D.T., Cardin A.D.,  
 RA Cortes P.E., Fenton J.W. II, Tulinsky A.;  
 RT "Crystallographic structure of human gamma-thrombin.";  
 RL J. Biol. Chem. 269:22000-22006(1994).

RN (10)  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE=97357286; PubMed=9214615;  
 RA van de Lochte A., Bode W., Huber R., Le Bonniec B.F., Stone S.R.,  
 RA Esmen C.T., Stubbs M.T.;  
 RT "The thrombin E192Q-38PT complex reveals gross structural  
 RT rearrangements: implications for the interaction with antithrombin  
 RT and thrombomodulin.";  
 RL EMO J. 16:2977-2984(1997).  
 RN (11)  
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 328-601.  
 RX MEDLINE=99162521; PubMed=10051558;  
 RA Guinto E.R., Caccia S., Rose T., Fueterer K., Waksman G., di Cera E.;  
 RT "Unexpected crucial role of residue 225 in serine proteases.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:1852-1857(1999).  
 RN (12)  
 RP VARIANT BARCELONA.  
 RX MEDLINE=87033739; PubMed=3771562;  
 RA Rablet M.-J., Furie B.C., Furie B.;  
 RT Molecular defect of prothrombin Barcelona. Substitution of cysteine  
 RT for arginine at residue 273.";  
 RL J. Biol. Chem. 261:15045-15048(1986).  
 RN (13)  
 RP VARIANT FRANKFURT.  
 RX MEDLINE=95313001; PubMed=7792730;  
 RA Degen S.J.F., McDowell S.A., Sparks L.M., Scharrer I.;  
 RT "Prothrombin Frankfurt: a dysfunctional prothrombin characterized by  
 RT substitution of Glu-466 by Ala.";  
 RL Thromb. Haemost. 73:203-209(1995).  
 RN (14)  
 RP VARIANTS HIMI-1 AND HIMI-2.  
 RX MEDLINE=93043342; PubMed=1421398;  
 RA Morishita E., Saito M., Kumabashiri I., Asakura H., Matsuda T.,  
 RA Yamaguchi K.;  
 RT "Prothrombin Himi: a compound heterozygote for two dysfunctional  
 RT prothrombin molecules (Met-337->Thr and Arg-388->His).";  
 RL Blood 80:2275-2280(1992).  
 RN (15)  
 RP VARIANT PADUA-1.  
 RX MEDLINE=95169896; PubMed=7865694;  
 RA James H.L., Kim D.J., Zheng D.-Q., Girolami A.;  
 RT "Prothrombin Padua I: incomplete activation due to an amino acid  
 RT substitution at a factor Xa cleavage site.";  
 RL Blood Coagul. Fibrinolysis 5:841-844(1994).  
 RN (16)  
 RP VARIANT QUICK-1.  
 RX MEDLINE=89207504; PubMed=3242619;  
 RA Henriksen R.A., Mann K.G.;  
 RT "Identification of the primary structural defect in the dysprothrombin  
 RT thrombin Quick I: substitution of cysteine for arginine-382.";  
 RL Biochemistry 27:9160-9165(1988).  
 RN (17)  
 RP VARIANT QUICK-2.  
 RX MEDLINE=89247386; PubMed=2719946;  
 RA Henriksen R.A., Mann K.G.;  
 RT "Substitution of valine for glycine-558 in the congenital dysprothrombin  
 RT thrombin Quick II alters primary substrate specificity.";

RL Biochemistry 28:2078-2082(1989).  
 RN (18)  
 RP VARIANT SALAKTA.  
 RX MEDLINE=92378975; PubMed=1354985;  
 RA Miyata T., Aruga R., Uneyama H., Beeasad A., Gullin M.-C.,  
 RA Iwanaga S.;  
 RT "Prothrombin Salakta: substitution of glutamic acid-466 by alanine  
 RT reduces the fibrinogen clotting activity and the esterase activity.";  
 RL Biochemistry 31:7457-7462(1992).  
 RN (19)  
 RP VARIANT TOKUSHIMA.  
 RX MEDLINE=87185407; PubMed=3567158;  
 RA Miyata T., Morita T., Inamoto T., Kawauchi S., Shirakami A.,  
 RA Iwanaga S.;  
 RT "Prothrombin Tokushima, a replacement of arginine-418 by tryptophan  
 RT that impairs the fibrinogen clotting activity of derived thrombin  
 RT Tokushima.";  
 RL Biochemistry 26:1117-1122(1987).  
 RN (20)  
 RP VARIANT TOKUSHIMA.  
 RX MEDLINE=87101511; PubMed=3801671;  
 RA Inamoto T., Shirakami A., Kawauchi S., Shigekiyo T., Saito S.,  
 RA Miyoshi K., Morita T., Iwanaga S.;  
 RT "Prothrombin Tokushima: characterization of dysfunctional thrombin  
 RT derived from a variant of human prothrombin.";  
 RL Blood 69:565-569(1987).  
 RN (21)  
 RP VARIANT TOKUSHIMA.  
 RX MEDLINE=92256895; PubMed=1349838;  
 RA Iwanaga S., Yoshimoto K., Shigekiyo T., Shirakami A., Saito S.,  
 RA Itakura M.;  
 RT "Detection of a single base substitution of the gene for prothrombin  
 RT Tokushima. The application of PCR-SSCP for the genetic and molecular  
 RT analysis of dysprothrombinemia.";  
 RL Int. J. Hematol. 55:93-100(1992).  
 RN (22)  
 RP VARIANT TYPE-3.  
 RX MEDLINE=83204687; PubMed=6405779;  
 RA Board P.G., Shaw D.C.;  
 RT "Determination of the amino acid substitution in human prothrombin  
 RT type 3 (157 Glu leads to Lys) and the localization of a third  
 RT thrombin cleavage site.";  
 RL Br. J. Haematol. 54:245-254(1983).  
 RN (23)  
 RP VARIANTS MET-163 AND THR-386.  
 RX MEDLINE=99318093; PubMed=10391209;  
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patel N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipschutz R., Daley G.G.;  
 RT "Characterization of single-nucleotide polymorphisms in coding regions  
 RT of human genes.";  
 RL Nat. Genet. 22:231-238(1999).  
 RN (24)  
 RP ERRATUM.  
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patel N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,

RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
 RA Lander E.S.;  
 RL Nat. Genet. 23:373-373(1999).  
 CC -1- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS  
 CC FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VIII, XIII,  
 CC AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C, I-gly, activates  
 CC -1- CATALYTIC ACTIVITY: preferential cleavage: Arg-1-Gly; activates  
 CC fibrinogen to fibrin and releases fibrinopeptide A and B.  
 CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN THE LIVER; FOUND IN PLASMA.  
 CC -1- PM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,  
 CC RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOXAL  
 CC ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES  
 CC ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY  
 CC CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION

Query Match 100.0%; Score 69; DB 1; Length 622;  
 Best Local Similarity 100.0%; Pred. No. 0.00031;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGPFV 12  
 |||||  
 DB 562 DACEGDSGPFV 573

RESULT 4  
 ID THRB\_BOVIN STANDARD; PRT; 625 AA.  
 AC P00735;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Prothrombin precursor (EC 3.4.21.5).  
 GN F2.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 CX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=88245190; PubMed=3379642;  
 RA Irwin D.M., Robertson K.A., Macgillivray R.T.A.;  
 RT "Structure and evolution of the bovine prothrombin gene";  
 RL J. Mol. Biol. 200:31-45(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RK MEDLINE=84203525; PubMed=6326805;  
 RA McGillivray R.T.A., Davie E.W.;  
 RT "Characterization of bovine prothrombin mRNA and its translation  
 RT product";  
 RL Biochemistry 23:1626-1634(1984).  
 RN [3]  
 RP SEQUENCE OF 44-625, DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.  
 RA Magnusson S., Sottrup-Jensen L., Petersen T.E., Cleaves H.;  
 RL (in) Hemker H.C., Veltkamp J.D. (eds.);  
 RL Boehrhaave symposium on prothrombin and related coagulation factors,

RL pp.25-46, Leiden University Press, Leiden (1975).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF ACTIVATION PEPTIDE 1.  
 RX MEDLINE=86296631; PubMed=3741841;  
 RA Park C.H., Tulinsky A.;  
 RT "three-dimensional structure of the kinkle sequence: structure of  
 RT prothrombin fragment 1";  
 RL Biochemistry 25:3577-3582(1986).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF ACTIVATION PEPTIDE 1.  
 RX MEDLINE=91311686; PubMed=1836869;  
 RA Sehadati T.-P., Tulinsky A., Skrzypczak-Jankun E., Park C.H.;  
 RT "Structure of bovine prothrombin fragment 1 refined at 2.25-A  
 RT resolution";  
 RL J. Mol. Biol. 220:481-494(1991).  
 RN [6]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF ACTIVATION PEPTIDE 1.  
 RX MEDLINE=92190185; PubMed=1547238;  
 RA Soriano-Garcia M., Padmanabhan K., de Vos A.M., Tulinsky A.;  
 RT "The Ca2+ ion and membrane binding structure of the Gla domain of Ca-  
 RT prothrombin fragment 1";  
 RL Biochemistry 31:2534-2566(1992).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE=92218459; PubMed=156020;  
 RA Martin P.D., Robertson W., Turk D., Huber R., Edwards B.F.P.;  
 RT "The structure of residues 7-16 of the A alpha-chain of human  
 RT fibrinogen bound to bovine thrombin at 2.3-A resolution";  
 RL J. Biol. Chem. 267:7911-7920(1992).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE=92389319; PubMed=1518046;  
 RA Brandstetter H., Turk D., Hoeffken H.W., Grosse D., Stuerzbecher J.,  
 RA Martin P.D., Edwards B.F.P., Bode W.;  
 RT "Refined 2.3 A X-ray crystal structure of bovine thrombin complexes  
 RT formed with the benzamidine and arginine-based thrombin inhibitors  
 RT NAPA, 4-TAPAP and WQPA. A starting point for improving  
 RT antithrombotics";  
 RL J. Mol. Biol. 226:1085-1089(1992).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF COMPLEX WITH ORNITHODORIN.  
 RX MEDLINE=97102783; PubMed=8947023;  
 RA van de Locht A., Stubbs M.T., Bode W., Friedrich T., Bollschweiler C.,  
 RA Hoeffken W., Huber R.;  
 RT "The ornithodorin-thrombin crystal structure, a key to the TAP  
 RT epitope?";  
 RL EMBO J. 15:6011-6017(1996).  
 RN [10]  
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH TRIPABIN.  
 RX MEDLINE=98004486; PubMed=9342325;  
 RA Fuentes-Prior P., Noeske-Jungblut C., Donner P., Schleuning W.D.,  
 RA Huber R., Bode W.;  
 RT "Structure of the thrombin complex with triabin, a lipocalin-like  
 RT exosite-binding inhibitor derived from a triatomine bug";  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:11045-11050(1997).  
 RN [11]  
 RP GENE STRUCTURE.



XX	MEDLINE=51358502; PubMed=1885621;
RA	Tsuji A., Torres-Rosado A., Arai T., le Beau M.M., Lemons R.S.,
RA	Kurahachi K.;
RT	"Hepsin, a cell membrane-associated protease. Characterization,
RT	tissue distribution, and gene localization.";
RL	J. Biol. Chem. 266:16946-16953(1991).
RN	14)
RP	CHARACTERIZATION.
RX	MEDLINE=93348237; PubMed=8346233;
RA	Torres-Rosado A., O'Shea K.S., Tsuji A., Chou S.H., Kurachi K.;
RT	"Hepsin, a putative cell-surface serine protease, is required for
RT	mamalian cell growth.";
RL	Proc. Natl. Acad. Sci. U.S.A. 90:7181-7187(1993).
CC	-1- FUNCTION: Plays an essential role in cell growth and maintenance
CC	of cell morphology.
CC	-1- SUBCELLULAR LOCATION: Type II membrane protein.
CC	-1- TISSUE SPECIFICITY: Present in most tissues, with the highest
CC	level in liver.
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMBL; M8930; AAA36013.1; -
DR	EMBL; X07732; CAA30558.1; -
DR	EMBL; X07702; CAA30058.1; -
DR	EMBL; BC025716; AA425716.1; -
DR	PIR; S00845; S00845.
DR	HSSP; P00763; IDPO.
DR	MEROPS; S01.224; -
DR	Genew; HGNC:5155; HPN.
DR	MM; 142440; -
DR	GO; GO:0005687; C:integral to plasma membrane; TAS.
DR	GO; GO:0008151; P:cell growth and/or maintenance; TAS.
DR	InterPro; IPR001314; Cytocytolipin.
DR	InterPro; IPR001254; Ser_Protease_Try.
DR	Pfam; PF00089; trypsin; 1.
DR	PRINTS; PR00722; CHRMOTRYPSIN.
DR	SMART; SM00020; TRYP_SP; 1.
DR	PROSITE; PSS0240; TRYPSIN_DOM; 1.
DR	PROSITE; PSS0134; TRYPSIN_HIS; 1.
DR	PROSITE; PSS0135; TRYPSIN_SER; 1.
KM	Hydrolase; Serine protease; Transmembrane; Signal-anchor.
FT	CHAIN 1 162
FT	SRINE PROTEASE HEPsin, NON-CATALYTIC
FT	CHAIN (POTENTIAL).
FT	CHAIN 163 417
FT	CATALYTIC CHAIN
FT	(POTENTIAL).
FT	DOMAIN 1 17
FT	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 18 44
FT	SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT	(POTENTIAL).
FT	DOMAIN 45 417
FT	EXTRACELLULAR (POTENTIAL).
FT	DOMAIN 163 417
FT	SERINE PROTEASE.

FT ACT SITE 203 203 CHANGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT SITE 257 257 CHANGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT SITE 353 353 CHANGE RELAY SYSTEM (BY SIMILARITY).  
 FT DISULFID 153 277 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 188 204 BY SIMILARITY.  
 FT DISULFID 322 338 BY SIMILARITY.  
 FT DISULFID 349 381 BY SIMILARITY.  
 FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 417 AA; 45011 MW; B2086FF6E1E551D7 CRC64;  
 Query Match 95.7%; Score 66; DB 1; Length 417;  
 Best Local Similarity 91.7%; Pred. No. 0.00067;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DACEGDSGGPFV 12  
 Db 347 DACQDGS GGPFV 358  
 RESULT 6  
 HEP5\_MOUSE STANDARD; PRT; 436 AA.  
 AC 035453; 09CW97;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Serine protease hepsin (EC 3.4.21.-).  
 GN HPN.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Liver;  
 RX MEDLINE=98058912; PubMed=9395459;  
 RA Vu T.-K.H., Liu R.W., Haasema C., Tomasek J.J., Howard E.W.;  
 RT "Identification and cloning of the membrane-associated serine  
 RT protease, hepsin, from mouse preimplantation embryos";  
 RL J. Biol. Chem. 272:31315-31320(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RX MEDLINE=99339944; PubMed=10411637;  
 RA Kawamura S., Kurachi S., Deyashiki Y., Kurachi K.;  
 RT "Complete nucleotide sequence, origin of isoform and functional  
 RT characterization of the mouse hepsin gene";  
 RL Eur. J. Biochem. 262:755-764(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC STRAIN=C57BL/6J; TISSUE=Kidney;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishii K., Kiyosawa H., Konno S., Imanaka I.,  
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Glass C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schiraldi L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Bocfell D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Kombeets P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Teyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohzuki S.,  
 RA Hayashizaki Y.;  
 RA "Functional annotation of a full-length mouse cDNA collection";  
 RL Nature 409:695-690(2001).  
 CC -1- FUNCTION: Plays an essential role in cell growth and maintenance  
 CC of cell morphology.  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1; Synonyms=1a;  
 CC IsoId=035453-1; Sequence=Displayed;  
 CC Note=Minor isoform;  
 CC IsoId=035453-2; Sequence=VSP\_007232;  
 CC Note=Major isoform;  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -1- CAUTION: Ref.3 sequence differs from that shown due to  
 CC frameshifts in positions 195, 191 and 233.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF030065; AAB84221.1; -;  
 DR EMBL: AK002694; BAB2289.2; ALT\_FRAME.  
 DR HSSP: P00763; IDPO.  
 DR MEROPS: S01.224; -;  
 DR MGd: MG1.1196620; Hpn.  
 DR InterPro: IPR001314; Cymotrypsin.  
 DR InterPro: IPR001254; Ser\_protease\_Try.  
 DR InterPro: IPR001190; Serl\_receptor.  
 DR Pfam: PF00069; trypsin\_1.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR SMART: SM00202; SPr\_1.  
 DR SMART: SM00202; TRYP\_SPr\_1.  
 DR PROSITE: PS00240; TRYPsin\_DOM\_1.  
 DR PROSITE: PS00134; TRYPsin\_HIS\_1.  
 DR PROSITE: PS00135; TRYPsin\_SER\_1.  
 KW Hydrolase; serine protease; Transmembrane; Signal-anchor;  
 KW Alternative splicing.  
 FT CHAIN 1 181 SERINE PROTEASE HEP5IN, NON-CATALYTIC  
 FT CHAIN 182 436 CHAIN (POTENTIAL).  
 FT CHAIN SERINE PROTEASE HEP5IN, CATALYTIC CHAIN

FT		(POTENTIAL).
ET	21	CYTOPLASMIC (POTENTIAL).
FT	37	SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).

FT	DOMAIN
FT 1	1.000000
FT 2	1.000000
FT 3	1.000000
FT 4	1.000000
FT 5	1.000000
FT 6	1.000000
FT 7	1.000000
FT 8	1.000000
FT 9	1.000000
FT 10	1.000000
FT 11	1.000000
FT 12	1.000000
FT 13	1.000000
FT 14	1.000000
FT 15	1.000000
FT 16	1.000000
FT 17	1.000000
FT 18	1.000000
FT 19	1.000000
FT 20	1.000000
FT 21	1.000000
FT 22	1.000000
FT 23	1.000000
FT 24	1.000000
FT 25	1.000000
FT 26	1.000000
FT 27	1.000000
FT 28	1.000000
FT 29	1.000000
FT 30	1.000000
FT 31	1.000000
FT 32	1.000000
FT 33	1.000000
FT 34	1.000000
FT 35	1.000000
FT 36	1.000000
FT 37	1.000000
FT 38	1.000000
FT 39	1.000000
FT 40	1.000000
FT 41	1.000000
FT 42	1.000000
FT 43	1.000000
FT 44	1.000000
FT 45	1.000000
FT 46	1.000000
FT 47	1.000000
FT 48	1.000000
FT 49	1.000000
FT 50	1.000000
FT 51	1.000000
FT 52	1.000000
FT 53	1.000000
FT 54	1.000000
FT 55	1.000000
FT 56	1.000000
FT 57	1.000000
FT 58	1.000000
FT 59	1.000000
FT 60	1.000000
FT 61	1.000000
FT 62	1.000000
FT 63	1.000000
FT 64	1.000000
FT 65	1.000000
FT 66	1.000000
FT 67	1.000000
FT 68	1.000000
FT 69	1.000000
FT 70	1.000000
FT 71	1.000000
FT 72	1.000000
FT 73	1.000000
FT 74	1.000000
FT 75	1.000000
FT 76	1.000000
FT 77	1.000000
FT 78	1.000000
FT 79	1.000000
FT 80	1.000000
FT 81	1.000000
FT 82	1.000000
FT 83	1.000000
FT 84	1.000000
FT 85	1.000000
FT 86	1.000000
FT 87	1.000000
FT 88	1.000000
FT 89	1.000000
FT 90	1.000000
FT 91	1.000000
FT 92	1.000000
FT 93	1.000000
FT 94	1.000000
FT 95	1.000000
FT 96	1.000000
FT 97	1.000000
FT 98	1.000000
FT 99	1.000000
FT 100	1.000000

FT	ACT_SITE
FT	ACT_SITE
FT	DISULEID
FT	DISULEID

FT DISULFID

CARBOHYD

100

FT : CONFLICT

## ET : CONFLICT

FT : CONFLICT

Best Local

```
Job: time : 5.
```

1 69 100.0 235 6 Q28731  
 2 69 100.0 235 13 Q91004  
 3 69 100.0 235 13 Q90387  
 4 69 100.0 239 13 Q91218  
 5 69 100.0 607 13 Q91001  
 6 69 100.0 608 13 Q9PTW7  
 7 66 95.7 234 13 Q90244  
 8 66 95.7 433 11 Q9CM97  
 9 66 95.7 1524 13 Q91674  
 10 63 92.8 420 13 Q90504  
 11 63 91.3 195 4 Q8J008  
 12 63 91.3 195 4 Q8J007  
 13 63 91.3 195 4 Q8J006  
 14 63 91.3 195 4 Q8J005  
 15 63 91.3 211 4 Q8J009  
 16 63 91.3 255 5 Q9NBC9  
 17 63 91.3 257 11 Q8BZ04  
 18 63 91.3 267 5 Q9BK47  
 19 63 91.3 356 5 Q45029  
 20 63 91.3 371 5 Q8MRX3  
 21 63 91.3 417 11 Q8B210  
 22 63 91.3 456 6 Q9TRP0  
 23 63 91.3 974 13 Q90WD8  
 24 63 91.3 1374 5 Q9VSU0  
 25 63 91.3 1449 5 Q9U112  
 26 63 91.3 1450 5 Q8IOB8  
 27 63 91.3 1462 5 Q9U113  
 28 63 91.3 2382 5 Q9B119  
 29 63 91.3 2409 5 Q960G6  
 30 63 91.3 2786 5 Q9VSU2  
 31 61 88.4 248 5 Q8IRF2  
 32 60 87.0 85 5 Q8MVL1  
 33 60 87.0 155 5 Q9YK4  
 34 60 87.0 187 5 Q45045  
 35 60 87.0 200 11 Q92406  
 36 60 87.0 234 11 Q9CV76  
 37 60 87.0 247 13 Q9W7Q5  
 38 60 87.0 250 5 Q9V514  
 39 60 87.0 252 5 Q76498  
 40 60 87.0 253 5 Q8SX24  
 41 60 87.0 253 5 Q8KX21  
 42 60 87.0 254 5 Q9XY70  
 43 60 87.0 254 5 Q76520  
 44 60 87.0 254 11 Q8CGR4  
 45 60 87.0 255 3 Q9Y7A9

# ALIGNMENTS

Q28731 oxytolagus  
 Q91004 gekko gekko  
 Q90387 cymops pyrr  
 Q91218 oncothymu  
 Q91001 gallus gall  
 Q9PTW7 struthio ca  
 Q90244 acipenser t  
 Q9CM97 mus musculu  
 Q91674 xenopus lae  
 Q90504 apteretus  
 Q8J008 homo sapien  
 Q8J007 homo sapien  
 Q8J006 homo sapien  
 Q8J005 homo sapien  
 Q8J009 homo sapien  
 Q9NBC9 glossina mo  
 Q8BZ04 mus musculu  
 Q9BK47 ludia foli  
 Q45029 drosophila  
 Q8MRX3 drosophila  
 Q8B210 mus musculu  
 Q9TRP0 canis famli  
 Q90WD8 bufo japoni  
 Q9VSU0 drosophila  
 Q9U112 drosophila  
 Q8IOB8 drosophila  
 Q9U113 drosophila  
 Q9B119 drosophila  
 Q960G6 drosophila  
 Q9VSU2 drosophila  
 Q8IRF2 drosophila  
 Q8MVL1 bolitena vi  
 Q9YK4 anopheles g  
 Q45045 scirpophaga  
 Q92406 mus musculu  
 Q9CV76 mus musculu  
 Q9W7Q5 paraliethy  
 Q9V514 drosophila  
 Q76498 diaprepes a  
 Q8SX24 drosophila  
 Q8KX21 drosophila  
 Q9XY70 rhyzopertha  
 Q76520 stromoxys ca  
 Q8CGR4 mus musculu  
 Q9Y7A9 metathidium

DE Thrombin (Fragment).  
 GN THROMBIN.  
 OS Oxytolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OK NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=9221913; PubMed=1557383;  
 RA Barfield D.K., MacGillivray R.T.A.;  
 RT "Partial characterization of vertebrate prothrombin cDNAs:  
 RT Amplification and sequence analysis of the B chain of thrombin from  
 RT nine different species.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).  
 DR EMBL; M61366; AA31477.1; -.  
 DR HSSP; P00734; 1UV5.  
 DR MEROPS; S01.217; -.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR003966; Prothrombin.  
 DR InterPro; IPR001254; Ser\_Protease\_Try.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR01905; PROTHROMBIN.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00136; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00133; TRYPSIN\_SER; 1.  
 KW Hydrolase; Protease; Serine protease.  
 FT NON\_TER 1  
 SQ SEQUENCE 235 AA; 27093 MW; 92FF3E4F938360E0 CRC64;  
 Query Match 100.0%; Score 69; DB 6; Length 235;  
 Best Local Similarity 100.0%; Pred. No. 0.00086;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DACBGSQGPFV 12  
 Db 176 DACBGSQGPFV 187  
 RESULT 2  
 Q91004  
 ID Q91004 PRELIMINARY; PRT; 235 AA.  
 AC Q91004;  
 DT 01-NOV-1996 (TEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
 DT 01-WAR-2003 (TEMBLrel. 23, Last annotation update)  
 DE Thrombin (Fragment).  
 GN THROMBIN.  
 OS Gecko gekko (Tokay gekko).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidosauria; Squamata; Scleroglossa; Gekkota; Gekkoniidae; Gekko.  
 OK NCBI\_TaxID=36310;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;

RX MEDLINE=92212913; PubMed=1557383;  
 RA Banfield D.K., MacGillivray R.T.A.;  
 RT "Partial characterization of vertebrate prothrombin cDNAs:  
 RT Amplification and sequence analysis of the B chain of thrombin from  
 RT nine different species."  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).  
 DR EMBL; M81392; AAA49309.1; -.  
 DR HSSP; P00734; 1B7X.  
 DR MEROPS; S01.217; -.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR003966; Prothrombin.  
 DR InterPro; IPR001254; Ser.protease\_Try.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR01505; PROTHROMBIN.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS0240; TRYPsin\_DOM; 1.  
 DR PROSITE; PS00134; TRYPsin\_HIS; 1.  
 DR PROSITE; PS00135; TRYPsin\_SER; 1.  
 KM Hydroxylase; Protease; Serine protease.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 235 AA; 26933 MW; 122A5C09F6F2276A CRC64;  
 Query Match 100.0%; Score 69; DB 13; Length 235;  
 Best Local Similarity 100.0%; Pred. No. 0.00086;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DACEGDSGGPFV 12  
 DB 175 DACEGDSGGPFV 186  
 RESULT 3  
 ID Q90387 PRELIMINARY; PRT; 235 AA.  
 AC Q90387;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Thrombin (Fragment).  
 GN THROMBIN.  
 OS Cynops pyrrhogaster (Japanese common newt).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Eukaryota; Batrachia; Caudata; Salamandridae; Cynops.  
 OC NCBI\_TaxID=8330;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=92212913; PubMed=1557383;  
 RA Banfield D.K., MacGillivray R.T.A.;  
 RT "Partial characterization of vertebrate prothrombin cDNAs:  
 RT Amplification and sequence analysis of the B chain of thrombin from  
 RT nine different species."  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).  
 DR EMBL; M81395; AAA49391.1; -.  
 DR HSSP; P00734; 1UVS.  
 DR MEROPS; S01.217; -.

DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR003966; Prothrombin.  
 DR InterPro; IPR001254; Ser.protease\_Try.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR01505; PROTHROMBIN.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS0240; TRYPsin\_DOM; 1.  
 DR PROSITE; PS00134; TRYPsin\_HIS; 1.  
 DR PROSITE; PS00135; TRYPsin\_SER; 1.  
 KM Hydroxylase; Protease; Serine protease.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 235 AA; 27272 MW; 49264D29A57A41F CRC64;  
 Query Match 100.0%; Score 69; DB 13; Length 235;  
 Best Local Similarity 100.0%; Pred. No. 0.00086;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DACEGDSGGPFV 12  
 DB 175 DACEGDSGGPFV 186  
 RESULT 4  
 ID Q91218 PRELIMINARY; PRT; 239 AA.  
 AC Q91218;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Thrombin (Fragment).  
 GN THROMBIN.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OC NCBI\_TaxID=8022;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=92212913; PubMed=1557383;  
 RA Banfield D.K., MacGillivray R.T.A.;  
 RT "Partial characterization of vertebrate prothrombin cDNAs:  
 RT Amplification and sequence analysis of the B chain of thrombin from  
 RT nine different species."  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).  
 DR EMBL; M81398; AAA49433.1; -.  
 DR HSSP; P00734; 1B7X.  
 DR MEROPS; S01.217; -.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR003966; Prothrombin.  
 DR InterPro; IPR001254; Ser.protease\_Try.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR01505; PROTHROMBIN.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS0240; TRYPsin\_DOM; 1.

DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolase; Protease; Serine protease.  
 FT NON TER 1  
 SQ SEQUENCE 239 AA; 27396 MW; F0F43F9A3205BF38 CRC64;

Query Match 100.0%; Score 69; DB 13; Length 239;  
 Best Local Similarity 100.0%; Pred. No. 0.00087;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGGPFV 12  
 |||||  
 Db 175 DACEGDSGGPFV 166

RESULT 5  
 ID Q91001 PRELIMINARY; PRT; 607 AA.  
 AC Q91001;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Thrombin.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=92212913; PubMed=1557383;  
 RA Banfield D.K., MacGillivray R.T.;  
 RT "Partial characterization of vertebrate prothrombin cDNAs:  
 RT amplification and sequence analysis of the B chain of thrombin from  
 RT nine different species."  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=94223694; PubMed=7513365;  
 RA Banfield D.K., Irwin D.M., Walz D.A., MacGillivray R.T.;  
 RT "Evolution of prothrombin: isolation and characterization of the cDNAs  
 RT encoding chicken and haggish prothrombin."  
 RL J. Mol. Evol. 38:177-187(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Banfield D.K.;  
 RL Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.  
 DR EMBL; M81391; AAA21619.1; -.  
 DR HSSP; P00734; IUVS.  
 DR MEROPS; S01.217; -.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR002383; GLA\_Blood.  
 DR InterPro; IPR000001; Kringle.

DR InterPro; IPR003966; Prothrombin.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR InterPro; IPR000294; VltK\_dep\_GLA.  
 DR Pfam; P00594; gla; 1.  
 DR Pfam; P00051; kringle; 2.  
 DR Pfam; P00089; trypsin; 1.  
 DR PRINTS; PRO0722; CHYMOTRYPSIN.  
 DR PRINTS; PRO0001; GLABLOOD.  
 DR PRINTS; PRO0018; KRINGLE.  
 DR PRINTS; PRO1905; PROTHROMBIN.  
 DR ProDom; PD000395; Kringle; 2.  
 DR SMART; SM00069; GLA; 1.  
 DR SMART; SM00130; KR; 2.  
 DR SMART; SM0020; Tryp\_spec; 1.  
 DR PROSITE; PS00011; GLU\_CARBOXYLATION; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 2.  
 DR PROSITE; PS50070; KRINGLE\_2; 2.  
 DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.  
 SQ SEQUENCE 607 AA; 69110 MW; 002F3606EA36270F CRC64;

Query Match 100.0%; Score 69; DB 13; Length 607;  
 Best Local Similarity 100.0%; Pred. No. 0.0022;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGGPFV 12  
 |||||  
 Db 548 DACEGDSGGPFV 559

RESULT 6  
 ID Q9PTW7 PRELIMINARY; PRT; 608 AA.  
 AC Q9PTW7;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Prothrombin.  
 GN Ostr.  
 OS Struthio camelus (Ostrich).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Archosauria; Aves; Palaeognathae; Struthioniformes; Struthionidae;  
 OC Struthio.  
 OX NCBI\_TaxID=8801;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=20579470; PubMed=11137455;  
 RA Frost C., Naude R., Oelofsen W., Muramoto K., Nagamura T., Ogawa T.;  
 RT "Purification and characterization of ostrich prothrombin."  
 RL Int. J. Biochem. Cell Biol. 32:1151-1159(2000).  
 CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.  
 DR EMBL; AB028871; BA869046.1; -.  
 DR HSSP; P00734; IUVS.  
 DR MEROPS; S01.217; -.  
 DR InterPro; IPR000001; Kringle.

DR InterPro: IPR001314; Chymotrypsin.  
DR InterPro: IPR002383; GLA\_blood.  
DR InterPro: IPR000001; Kringler.  
DR InterPro: IPR003666; Prothrombin.  
DR InterPro: IPR001254; Ser.protease\_Try.  
DR InterPro: IPR000294; VAtk\_dep\_GLA.  
DR Pfam: PF00594; gla; 1.  
DR Pfam: PF00051; Kringler; 2.  
DR Pfam: PF00089; trypsin; 1.  
DR PRINTS: PR00722; CHYMOTRYPSIN.  
DR PRINTS: PR00001; GLABLOOD.  
DR PRINTS: PR00018; KRINGLE.  
DR PRINTS: PR01505; PROTHROMBIN.  
DR ProDom: PD000385; Kringler; 2.  
DR SMART: SMO0069; GLA; 1.  
DR SMART: SMO0130; KR; 2.  
DR SMART: SMO0020; Tryp\_Spec; 1.  
DR PROSITE: PS00011; GLU CARBOXYLATION; 1.  
DR PROSITE: PS00021; KRINGLE\_1; 2.  
DR PROSITE: PS50070; KRINGLE\_2; 2.  
DR PROSITE: PS50240; TRYPsin\_DOM; 1.  
DR PROSITE: PS00134; TRYPsin\_HIS; 1.  
DR PROSITE: PS00135; TRYPsin\_SER; 1.  
KW Glycoprotein; Hydrolase; Kringler; Protease; Serine protease.  
SQ SEQUENCE 608 AA; 69392 MW; 11B974B9AE54EA2 CRC64;

Query Match 100.0%; Score 69; DB 13; Length 608;  
Best Local Similarity 100.0%; Pred. No. 0.0022;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DACEGDSGGPFV 12  
|||||  
Db 548 DACEGDSGGPFV 559

RESULT 7  
ID Q90244 PRELIMINARY; PRT; 234 AA.

AC Q90244; 01-JUN-2001 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Thrombin (Fragment).  
GN THROMBIN.  
OS *Acipenser transmontanus* (White sturgeon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;  
OC Acipenser.  
OX NCBI\_TaxID=7904;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=92212913; PubMed=1557383;  
RA Banfield D.K., MacGillivray R.T.A.;  
RT "Partial characterization of vertebrate prothrombin cDNAs:  
RT Amplification and sequence analysis of the B chain of thrombin from  
RT nine different species.";

RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).  
DR EMBL: M61399; AAA48514.1; -.  
DR HSSP: P00734; ZHNT.  
DR MEROPS: S01.217; -.  
DR InterPro: IPR001314; Chymotrypsin.  
DR InterPro: IPR003666; Prothrombin.  
DR InterPro: IPR001254; Ser.protease\_Try.  
DR Pfam: PF00089; trypsin; 1.  
DR Pfam: PF00051; Kringler; 2.  
DR PRINTS: PR00722; CHYMOTRYPSIN.  
DR PRINTS: PR01505; PROTHROMBIN.  
DR SMART: SMO0020; Tryp\_Spec; 1.  
DR PROSITE: PS50240; TRYPsin\_DOM; 1.  
DR PROSITE: PS00134; TRYPsin\_HIS; 1.  
DR PROSITE: PS00135; TRYPsin\_SER; 1.  
KW Hydrolase; Protease; Serine protease.  
FT NON\_TER 1  
SQ SEQUENCE 234 AA; 26846 MW; 43C558D618E0585 CRC64;

Query Match 95.7%; Score 66; DB 13; Length 234;  
Best Local Similarity 91.7%; Pred. No. 0.0027;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DACEGDSGGPFV 12  
|||||  
Db 175 DACEGDSGGPFV 186

RESULT 8

Qy 1 DACEGDSGGPFV 12  
|||||  
Db 175 DACEGDSGGPFV 186

AC Q9Cw97; PRELIMINARY; PRT; 435 AA.  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Adult male kidney cDNA, RIKEN full-length enriched library,  
DE clone:0610030A17 product:hepsin, full insert sequence.  
OS *Mus musculus* (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Kidney;  
RA Adechi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
RA Arawaka T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
RA Hamaoka T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,  
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
RA Kawai J., Kojima Y., Komio H., Kouda M., Koyu S., Kurihara C.,  
RA Matsuyama T., Miyazaki A., Nishikawa K., Nomura K., Numazaki R., Ohno M.,  
RA Okazaki Y., Okido T., Owa C., Saito R., Sakai C., Sakai K.,  
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
RA Muramatsu M., Hayashizaki Y.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Kidney;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium;  
 RA The RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Kidney;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Kidney;  
 RX MEDLINE=99279253; PubMed=10349636;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Kidney;  
 RX MEDLINE=20499374; PubMed=11042159;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Kidney;  
 RX MEDLINE=20350913; PubMed=11076861;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
 Kono H., Akiyama J., Nishi K., Katsuna T., Tashiro H., Itoh M.,  
 Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,  
 Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
 Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multiplexed cDNA sequences.";  
 RL Genome Res. 10:1757-1771(2000).  
 DR EMBL: AK002694; BAB22289.2; -  
 SQ SEQUENCE 435 AA; 45944 MW; 019B2A9D3EBEF40 CRC64;

Query Match 95.7%; Score 66; DB 11; Length 435;  
 Best Local Similarity 91.7%; Pred. No. 0.0051;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGGPFV 12  
 |||:|||||||  
 Db 365 DACEGDSGGPFV 376

Q91674  
 ID Q91674 PRELIMINARY; PRT; 1524 AA.  
 AC Q91674;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Polypeptide.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 OC Xenoidea; Xenopus.  
 OC NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99432219; PubMed=10500163;  
 RA Lindsay L.L., Yang J.C., Hedrick J.L.;  
 RT "Ovocytase, a Xenopus laevis egg extracellular protease, is  
 RT translated as part of an unusual polypeptide.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:11253-11258(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Yang J.C., Lindsay L.L., Hedrick J.L.;  
 RT "cDNA Cloning of Ovocytochrome, a Chymotrypsin-like Protease Released  
 RT From Xenopus laevis Eggs at Fertilization.";  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -1- SIMILARITY: CONTAINS 4 CUB DOMAINS.  
 DR EMBL: U81290; AAC24717.1; -  
 DR HSP70; P00763; IDPO.  
 DR MEROPS; S01.022; -  
 DR MEROPS; S01.245; -  
 DR InterPro: IPR001314; Chymotrypsin.  
 DR InterPro: IPR000859; CUB\_domain.  
 DR InterPro: IPR001254; Ser\_protease\_Try.  
 DR Pfam: PF00431; CUB; 5.  
 DR Pfam: PF00089; trypsin; 3.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR SMART: SM00042; CUB; 4.  
 DR SMART: SM00020; TRY-SPC; 3.  
 DR PROSITE: PS0180; CUB; 5.  
 DR PROSITE: PS0240; TRYPsin\_DOM; 3.  
 DR PROSITE: PS00134; TRYPsin\_HIS; 3.  
 DR PROSITE: PS00135; TRYPsin\_SER; 3.  
 KW Hydrolase; Protease; Serine protease.  
 FT CHAIN 57 308 SERINE PROTEASE.  
 FT CHAIN 1295 1524 SERINE PROTEASE.  
 FT CHAIN 1295 1524 OVOCHYMASE.  
 SQ SEQUENCE 1524 AA; 167566 MW; 32FE4212EF37269 CRC64;

Query Match 95.7%; Score 66; DB 13; Length 1524;  
 Best Local Similarity 91.7%; Pred. No. 0.018;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DACEGDSGGPFV 12  
 |||:|||||||  
 Db 241 DACEGDSGGPFV 252

RESULT 9

RESULT 10

Q0504 ID Q0504 PRELIMINARY; PRT; 420 AA.

AC Q0504; 01-NOV-1996 (TEMBLrel. 01, Created)

DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)

DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)

DE Thrombin.

OS *Eptatretus stoullii* (Pacific hagfish).

OC Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniiformes;

OC Myxiniidae; Eptatretinae; Eptatretus.

OX NCBI\_TaxID=7769;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RX MEDLINE=92212913; PubMed=1597383;

RA Banfield D.K., MacGillivray R.T.;

RT "Partial characterization of vertebrate prothrombin cDNAs: amplification and sequence analysis of the B chain of thrombin from nine different species."

RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RX MEDLINE=94223694; PubMed=7513365;

RA Banfield D.K., Irwin D.M., Walz D.A., MacGillivray R.T.;

RT "Evolution of prothrombin: isolation and characterization of the cDNAs encoding chicken and hagfish prothrombin."

RL J. Mol. Evol. 38:177-187(1994).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RL Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.

CC -1 SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

DR EMBL; M81393; AAA21620.1; -.

DR HSSP; P00734; IUVS.

DR MEROPS; S01.217; -.

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR003966; Prothrombin.

DR InterPro; IPR001254; Ser\_protease\_Try.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PRO0016; KRINGLE.

DR PRINTS; PRO1505; PROTHROMBIN.

DR ProDom; PD000395; Kringle; 1.

DR SMART; SM00130; KR; 1.

DR SMART; SM00020; Tryp\_Spc; 1.

DR PROSITE; PS00021; KRINGLE\_1; 1.

DR PROSITE; PS00070; KRINGLE\_2; 1.

DR PROSITE; PS00240; TRYPSIN\_DOM; 1.

DR PROSITE; PS00134; TRYPSIN\_HIS; 1.

DR PROSITE; PS00135; TRYPSIN\_SER; 1.

KW Hydrolase; Kringle; Protease; Serine protease.

SQ SEQUENCE 420 AA; 47888 MW; 64522AA21A57B67A CRC64;

Query Match 92.8%; Score 64; DB 13; Length 420;

Best Local Similarity 91.7%; Pred. No. 0.011;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DACEGDSGGPFV 12

DB 359 DPCEGDSGGPFV 370

Search completed: February 11, 2004, 14:56:04

Job time : 22.5161 secs

OK protein - protein search, using SW model

Run on: February 11, 2004, 14:35:52 ; Search time 49.7097 Seconds

(without alignments)  
73.441 Million cell updates/sec

Title: US-10-050-611-3

Perfect score: 131

Sequence: 1 AGYKPDGKRGDAGDSGGPRV 23

Scoring table: ELOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,

ALIGNMENTS

Result No.	Score	Match	Length	DB	ID	Description
1	131	100.0	23	20	AAW63414	Cell growth/adhesi
2	131	100.0	23	21	AAI12893	Nerve tissue regen
3	131	100.0	23	22	AAI70363	Human thrombin rec
4	131	100.0	23	23	AAE22563	Human thrombin big
5	131	100.0	23	23	AAE20159	Human thrombin pep
6	131	100.0	23	23	AAI50858	Thrombin-derived p
7	131	100.0	116	20	AAW91115	Human zeta 2 preth
8	131	100.0	259	18	AAW11545	Human thrombin Asn
9	131	100.0	259	24	APP60563	Human thrombin var
10	131	100.0	259	24	APP60565	Human thrombin var
11	131	100.0	295	16	AAI74775	Wild-type thrombin
12	131	100.0	295	16	AAI74776	Mutant thrombin K5
13	131	100.0	295	16	AAI74777	Mutant thrombin E2
14	131	100.0	295	16	AAI74778	Mutant thrombin E2
15	131	100.0	295	16	AAI74779	Mutant thrombin E2
16	131	100.0	295	16	AAI74780	Mutant thrombin E2
17	131	100.0	295	16	AAI76033	Mutant thrombin E2
18	131	100.0	295	16	AAI76034	Mutant thrombin R2
19	131	100.0	295	16	AAI76035	Mutant thrombin R2
20	131	100.0	295	16	AAI76036	Mutant thrombin R2
21	131	100.0	295	16	AAI76037	Mutant thrombin W5
22	131	100.0	295	16	AAI76038	Mutant thrombin K5
23	131	100.0	295	16	AAI76039	Mutant thrombin W5
24	131	100.0	295	16	AAI76040	Mutant thrombin W5
25	131	100.0	295	18	AAW22882	Human mature throm
26	131	100.0	295	21	AAI09633	Amino acid sequenc
27	131	100.0	295	24	AAI09632	Human thrombin var
28	131	100.0	295	24	AAI09634	Human thrombin var
29	131	100.0	308	20	AAW91099	Human prethrombin
30	131	100.0	376	14	AAI41787	CD4/Thrombin fusio
31	131	100.0	376	20	AAI42789	Human CD4/thrombin
32	131	100.0	376	23	AAU10703	Human CD4-thrombin
33	131	100.0	579	14	AAI35763	Prothrombin (Pr).
34	131	100.0	579	18	AAI11546	Human prothrombin
35	131	100.0	579	18	AAI11544	Human prothrombin
36	131	100.0	579	20	AAW91088	Human prothrombin
37	131	100.0	615	14	AAI38741	Human prothrombin
38	131	100.0	615	17	AAI38741	Human prothrombin
39	131	100.0	615	17	AAI38741	Human prothrombin
40	131	100.0	622	18	AAI11543	Human prothrombin
41	131	100.0	622	20	AAI45566	Human prothrombin
42	131	100.0	622	24	AAI45566	Human prothrombin
43	124	94.7	111	20	AAW91113	Human zeta 2 preth
44	124	94.7	308	20	AAW91107	Bovine prothrombin
45	124	94.7	582	20	AAW91106	Bovine prothrombin

RESULT 1  
 ID AAM63414 standard; peptide; 23 AA.  
 AC AAM63414;  
 XX  
 DT 26-FEB-1999 (first entry)  
 XX  
 DE Cell growth/adhesion promoting peptide #1.  
 XX  
 KW Cell growth; adhesion; promotion; medical treatment; injury;  
 KW biotissue; bone reinforcement; nerve regeneration; HRP resin.  
 XX  
 OS Synthetic.  
 XX  
 PN JP10316581-A.  
 XX  
 PD 02-DEC-1998.  
 XX  
 PF 15-MAY-1997; 97JP-0140685.  
 XX  
 PR 15-MAY-1997; 97JP-0140685.  
 XX  
 PA (KURS ) KURARAY CO LTD.  
 XX  
 DR WPI; 1999-076400/07.  
 XX  
 PT Material for medical treatment comprises new peptide - used for  
 PT covering injuries, promoting adhesion of bio-tissues, bone  
 PT reinforcing and nerve regeneration  
 XX  
 PS Claim 1; Page 12; 14pp; Japanese.  
 XX  
 CC The present invention describes a material for medical treatment which  
 CC comprises one or more peptides of the formula XADBGJLMPQY, or their  
 CC salts, immobilised on a substrate: where X = H, CH3CO or CH3COOlys;  
 CC A = Ser or Thr; D = Ile, Val or Leu; E = Lys or Arg; G = Ile, Val or  
 CC Leu; J = Gly or Ala; L = Ile, Val or Leu; M = Gly or Ala; Q = Gly, Ala  
 CC or Gly-Lys-Lys-Gly; Y = OH or NH2. Also described is an agent for cell  
 CC growth promotion and/or cell adhesion promotion containing the above  
 CC peptide or its salt as the active component. The peptide and its salt  
 CC can be used for covering injuries, promoting adhesion of biotissues,  
 CC bone reinforcing and nerve regeneration. The present sequence represents  
 CC a specifically claimed peptide of the present invention.  
 XX  
 SQ Sequence 23 AA;  
 Query Match 100.0%; Score 131; DB 20; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 3,4e-08;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGYKPEDEGRGDACEGDSGPFV 23  
 |||  
 Db 1 AGYKPEDEGRGDACEGDSGPFV 23

AAB12893  
 ID AAB12893 standard; peptide; 23 AA.  
 AC AAB12893;  
 XX  
 DT 02-NOV-2000 (first entry)  
 XX  
 DE Nerve tissue regenerative peptide SEQ ID #8.  
 XX  
 KW Nerve regeneration; nerve cell proliferation; axon extension; treatment;  
 KW central nervous system disorder; peripheral nervous system disorder;  
 KW spinal disorder; head injury; cerebrovascular disorder.  
 XX  
 OS Synthetic.  
 XX  
 PN JP2000143531-A.  
 XX  
 PD 23-MAY-2000.  
 XX  
 PF 11-AUG-1999; 99JP-0227108.  
 XX  
 PR 09-SEP-1998; 98JP-0270498.  
 XX  
 PA (KURS ) KURARAY CO LTD.  
 PA (NISH/) NISHIMURA Y.  
 PA (SUZU/) SUZUKI Y.  
 PA (TANI/) TANIHARA M.  
 XX  
 DR WPI; 2000-415772/36.  
 XX  
 PT New nerve regeneration material -  
 PT  
 XX  
 PS Claim 2; Page 5; 17pp; Japanese.  
 XX  
 CC This invention relates to a new nerve regenerative material which  
 CC contains a peptide immobilised to a base which consists of a  
 CC polyaccharide gel, such as alginic acid. Sequences AAB12886-B12899  
 CC represent examples of the peptides used in the nerve regeneration  
 CC material. The peptide containing material causes nerve cell  
 CC proliferation and also causes axonal extension. The material can be used  
 CC for the treatment of central or peripheral nervous system disorders,  
 CC spinal disorders, head injury or cerebrovascular disorders.  
 XX  
 SQ Sequence 23 AA;  
 Query Match 100.0%; Score 131; DB 21; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 3,4e-08;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGYKPEDEGRGDACEGDSGPFV 23  
 |||  
 Db 1 AGYKPEDEGRGDACEGDSGPFV 23

RESULT 2

RESULT 3  
 AAB70363  
 ID AAB70363 standard; peptide; 23 AA.







DR WPI; 1997-065455/06.  
 XX Prothrombin mutants with reduced clotting activity - useful as  
 PT antagonists of thrombin inhibitors or for anticoagulant therapy  
 XX  
 PS Example 3; Page -; 73pp; German.  
 XX  
 CC Prothrombin mutants having one or more changes in amino acid sequence  
 CC compared with the natural protein and having 0-10% (preferably 0-0.25%)  
 CC of the activity of the natural protein are claimed, provided that the  
 CC changes in amino acid sequence do not affect the capacity of the  
 CC mutants to bind to specific ligands and receptors. The mutants have  
 CC greatly reduced clotting activity and are useful as antagonists of  
 CC thrombin inhibitors such as hirudin, heparin and anti-thrombin III.  
 CC The mutations may also result in changes to the in vivo half-life  
 CC of prothrombin. The half-life may be reduced to less than 10 minutes  
 CC or the mutant prothrombin may have an extended half-life of more than  
 CC 1 hour, making it useful as an anticoagulant and to inhibit side-  
 CC effects of anti-coagulant treatment. They are converted to inactive  
 CC thrombin and are able to compete with native, active thrombin for  
 CC binding to receptors. The present sequence represents the thrombin  
 CC mutant which is derived by trypsin cleavage of a specifically  
 CC claimed human prothrombin mutant in which Asp at position 419 is  
 CC changed to Asn. The thrombin Asn99 mutant was found to have only  
 CC 0.24% of the activity of wild-type thrombin on a chromogenic  
 CC substrate.  
 CC (Note: This sequence does not appear in the specification and has  
 CC been produced by modifying the wild-type sequence of human  
 CC prothrombin which appears in figure 1).  
 XX  
 SQ Sequence 259 AA;  
 XX  
 Query Match 100.0%; Score 131; DB 18; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-07;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGYKPDGKRGDACEGDSGSPFV 23  
 ||||||||||||||||||  
 DB 188 AGYKPDGKRGDACEGDSGSPFV 210

RESULT 9  
 ABP60563  
 ID ABP60563 standard; protein; 259 AA.  
 XX  
 AC ABP60563;  
 XX  
 DT 28-MAR-2003 (first entry)  
 XX  
 DE Human thrombin variant W215A B-chain.  
 XX  
 KW Human; thrombin; W215A; anticoagulant; prothrombin; antithrombotic;  
 KW thrombus; protein C activation.  
 XX  
 OS Homo sapiens.  
 KW  
 XX  
 FH Key Location/Qualifiers

FT Misc-difference 229  
 FT /note= "Wild-type Trp substituted by Ala"  
 XX  
 XX W02002100337-A2.  
 XX  
 PD 19-DEC-2002.  
 XX  
 XX 07-JUN-2002; 2002WC-US16211.  
 XX  
 XX 08-JUN-2001; 2001US-297059P.  
 XX  
 XX (UYEK-) UNIV EMORY.  
 PA  
 XX Gruber A, Hanson SR, Di Cera E;  
 PT WPI; 2003-156907/15.  
 DR  
 XX  
 PT New variant thrombin, useful as an antithrombotic agent for inhibiting  
 PT the formation of a thrombus, for determining the level of protein C  
 PT activation in a blood sample, or for determining the thrombogenic  
 PT potential of a patient -  
 XX  
 XX Claim 15; Fig 2; 95pp; English.  
 PS  
 CC The invention relates to a novel variant human thrombin. The thrombin  
 CC variant of the invention has anticoagulant activity. The variant thrombin  
 CC or prothrombin is useful as an antithrombotic agent for inhibiting the  
 CC formation of a thrombus. The variant thrombin is also useful for  
 CC determining the level of protein C activation in a blood sample or the  
 CC thrombogenic potential of a patient. The present sequence represents the  
 CC B-chain of the thrombin variant W215A.  
 XX  
 SQ Sequence 259 AA;  
 XX  
 Query Match 100.0%; Score 131; DB 24; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-07;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGYKPDGKRGDACEGDSGSPFV 23  
 ||||||||||||||||||  
 DB 188 AGYKPDGKRGDACEGDSGSPFV 210

RESULT 10  
 ABP60565  
 ID ABP60565 standard; protein; 259 AA.  
 XX  
 AC ABP60565;  
 XX  
 DT 28-MAR-2003 (first entry)  
 XX  
 DE Human thrombin variant W215A/E217A B-chain.  
 XX  
 KW Human; thrombin; W215A/E217A; anticoagulant; prothrombin; antithrombotic;  
 KW thrombus; protein C activation.  
 XX  
 OS Homo sapiens.  
 KW  
 XX

XX Key Location/Qualifiers  
 FH Thrombin; oligonucleotide-directed mutagenesis; Procoagulant;  
 FT Misc-difference 227 /note= "Wild-type Trp substituted by Ala"  
 FT Misc-difference 229 /note= "Wild-type Glu substituted by Ala"  
 FT /note= "Wild-type Glu substituted by Ala"  
 XX W02002100337-A2.  
 PN 19-DEC-2002.  
 XX 07-JUN-2002; 2002WO-US18211.  
 PF 08-JUN-2001; 2001US-297089P.  
 PR (UYEM-) UNIV EMORY.  
 XX Gruber A, Hanson SR, Di Cera E;  
 PI WPI: 2003-156907/15.  
 DR N-PSDB; AB225355.  
 XX New variant thrombin, useful as an antithrombotic agent for inhibiting  
 PT the formation of a thrombus, for determining the level of protein C  
 PT activation in a blood sample, or for determining the thrombogenic  
 PT potential of a patient -  
 XX Claim 2; Fig 4; 95pp; English.  
 PS The invention relates to a novel variant human thrombin. The thrombin  
 CC variant of the invention has anticoagulant activity. The variant thrombin  
 CC or prothrombin is useful as an antithrombotic agent for inhibiting the  
 CC formation of a thrombus. The variant thrombin is also useful for  
 CC determining the level of protein C activation in a blood sample or the  
 CC thrombogenic potential of a patient. The present sequence represents the  
 CC B-chain of the thrombin variant W215A/E217A (WE).  
 XX Sequence 259 AA;  
 SQ  
 Query Match 100.0%; Score 131; DB 24; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-07;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGKPPDEGKRGDACGDSGGPFV 23  
 ||||||||||||||||||  
 DB 188 AGKPPDEGKRGDACGDSGGPFV 210  
 RESULT 11  
 AAR74775  
 ID AAR74775 standard; Protein; 295 AA.  
 XX AAR74775;  
 AC  
 XX 25-MAR-2003 (updated)  
 DT 04-NOV-1995 (first entry)  
 XX

DE Wild-type thrombin.  
 XX Thrombin; oligonucleotide-directed mutagenesis; Procoagulant;  
 KW anticoagulant; protein engineering; ss.  
 XX Homo sapiens.  
 OS  
 FH Key Location/Qualifiers  
 FT Protein 37..295  
 FT /note= "mature protein"  
 XX W09513365-A2.  
 PN 18-MAY-1995.  
 PD 14-NOV-1994; 94WO-US13104.  
 PF 10-JUN-1994; 94US-0258038.  
 PR 12-NOV-1993; 93US-0152657.  
 XX (GILE-) GILEAD SCI.  
 PA  
 XX Gibbs CS, Leung LLK, Tsiang M;  
 PI WPI: 1995-194103/25.  
 DR N-PSDB; AA092455.  
 XX Thrombin derives with segregated pro- and anticoagulant activities -  
 PT useful for treating thrombotic disorders but also diagnosis,  
 PT treatment of tumours, etc.  
 XX Disclosure; Fig 1; 78pp; English.  
 PS The sequence represents wild-type (reference) thrombin. Mutants  
 CC of this sequence (AAR74776-80 and AAR76033-41) have at least 80%  
 CC homology with thrombin, and are capable of protein-C activation  
 CC without significant fibrinogen clotting activity, and vice versa  
 CC (specifically have a ratio of protein-C activity to fibrinogen  
 CC clotting activity of less than 0.5 or greater than 2 compared to  
 CC thrombin). The mutant thrombin sequences, produced in recombinant  
 CC cell culture or by in vitro methods, and are used to treat  
 CC thrombotic conditions, particularly during cardiac bypass surgery  
 CC and in cases of septic shock.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX Sequence 295 AA;  
 SQ  
 Query Match 100.0%; Score 131; DB 16; Length 295;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-07;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGKPPDEGKRGDACGDSGGPFV 23  
 ||||||||||||||||||  
 DB 224 AGKPPDEGKRGDACGDSGGPFV 246  
 RESULT 12

AA74776  
 ID AA74776 standard; Protein; 295 AA.  
 XX  
 AC AA74776;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 04-NOV-1995 (first entry)  
 XX  
 DE Mutant thrombin K52A, R233A.  
 XX  
 KW Thrombin; oligonucleotide-directed mutagenesis; procoagulant;  
 KW anticoagulant; protein engineering; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 88 /note= "Lys in wild-type"  
 FT Misc-difference 269 /note= "Arg in wild-type"  
 FT Protein 37..295 /note= "mature protein"  
 FT  
 XX  
 PN W09513385-A2.  
 XX  
 PD 18-MAY-1995.  
 XX  
 PF 14-NOV-1994; 94WO-US13104.  
 XX  
 PR 10-JUN-1994; 94US-0258038.  
 PR 12-NOV-1993; 93US-0152657.  
 XX  
 XX  
 PA (GILE-) GILEAD SCI.  
 XX  
 PI Gibbs CS, Leung LK, Tsang M;  
 XX  
 DR WPI; 1995-194103/25.  
 XX  
 PT Thrombin derives with segregated pro- and anticoagulant activities -  
 PT useful for treating thrombotic disorders but also diagnosis,  
 PT treatment of tumours, etc.  
 XX  
 PS Claim 22; Page 63/3; 78pp; English.  
 XX  
 CC The mutant thrombin sequence, generated by oligonucleotide-directed  
 CC mutagenesis, has at least 80% homology with thrombin, and is  
 CC capable of protein-C activation without significant fibrinogen  
 CC clotting activity, and vice versa (specifically, it has a ratio  
 CC of protein-C activity to fibrinogen clotting activity of less than  
 CC 0.5 or greater than 2 compared to thrombin). The mutant thrombin  
 CC is produced in recombinant cell culture or by in vitro methods,  
 CC and is used to treat thrombotic conditions, particularly during  
 CC cardiac bypass surgery and in cases of septic shock.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 295 AA;

Query Match 100.0%; Score 131; DB 16; Length 295;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-07;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGYKPEGRGRCGDSGSPV 23  
 |||  
 Db 224 AGYKPEGRGRCGDSGSPV 246  
 RESULT 13  
 AA74777  
 ID AA74777 standard; Protein; 295 AA.  
 XX  
 AC AA74777;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 04-NOV-1995 (first entry)  
 XX  
 DE Mutant thrombin E229D.  
 XX  
 KW Thrombin; oligonucleotide-directed mutagenesis; procoagulant;  
 KW anticoagulant; protein engineering; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 265 /note= "Glu in wild-type"  
 FT Protein 37..295 /note= "mature protein"  
 FT  
 XX  
 PN W09513385-A2.  
 XX  
 PD 18-MAY-1995.  
 XX  
 PF 14-NOV-1994; 94WO-US13104.  
 XX  
 PR 10-JUN-1994; 94US-0258038.  
 PR 12-NOV-1993; 93US-0152657.  
 XX  
 XX  
 PA (GILE-) GILEAD SCI.  
 XX  
 PI Gibbs CS, Leung LK, Tsang M;  
 XX  
 DR WPI; 1995-194103/25.  
 XX  
 PT Thrombin derives with segregated pro- and anticoagulant activities -  
 PT useful for treating thrombotic disorders but also diagnosis,  
 PT treatment of tumours, etc.  
 XX  
 PS Claim 22; Page 63/3; 78pp; English.  
 XX  
 CC The mutant thrombin sequence, generated by oligonucleotide-directed  
 CC mutagenesis, has at least 80% homology with thrombin, and is  
 CC capable of protein-C activation without significant fibrinogen  
 CC clotting activity, and vice versa (specifically, it has a ratio  
 CC of protein-C activity to fibrinogen clotting activity of less than

CC 0.5 or greater than 2 compared to thrombin). The mutant thrombin  
 CC is produced in recombinant cell culture or by in vitro methods,  
 CC and is used to treat thrombotic conditions, particularly during  
 CC cardiac bypass surgery and in cases of septic shock.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 295 AA;  
 Query Match 100.0%; Score 131; DB 16; Length 295;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-07;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGXPDEGKRGDACEGDSGSPFV 23  
 DB 224 AGXPDEGKRGDACEGDSGSPFV 246  
 RESULT 14  
 ID AAR74778 standard; Protein; 295 AA.  
 XX  
 AC AAR74778;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 04-NOV-1995 (first entry)  
 XX  
 DE Mutant thrombin E229F.  
 XX  
 KW Thrombin; oligonucleotide-directed mutagenesis; procoagulant;  
 KW anticoagulant; protein engineering; ss.  
 XX  
 CS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 265 /note= "Glu in wild-type"  
 FT Protein 37..295  
 FT /note= "mature protein"  
 XX  
 PN W09513385-A2.  
 XX  
 PD 18-MAY-1995.  
 XX  
 PF 14-NOV-1994; 94WO-US13104.  
 XX  
 PR 10-JUN-1994; 94US-0258038.  
 PR 12-NOV-1993; 93US-0152657.  
 XX  
 PA (GILE-) GILEAD SCI.  
 XX  
 PI Gibbs CS, Leung LK, Tsiang M;  
 XX  
 DR WPI; 1995-194103/25.  
 XX  
 FT Thrombin derivs with segregated pro- and anticoagulant activities -  
 FT useful for treating thrombotic disorders but also diagnosis,  
 FT treatment of tumours, etc.

XX Claim 22; Page 63/3; 78pp; English.  
 PS The mutant thrombin sequence, generated by oligonucleotide-directed  
 XX mutagenesis, has at least 80% homology with thrombin, and is  
 CC capable of protein-C activation without significant fibrinogen  
 CC clotting activity, and vice versa (specifically, it has a ratio  
 CC of protein-C activity to fibrinogen clotting activity of less than  
 CC 0.5 or greater than 2 compared to thrombin). The mutant thrombin  
 CC is produced in recombinant cell culture or by in vitro methods,  
 CC and is used to treat thrombotic conditions, particularly during  
 CC cardiac bypass surgery and in cases of septic shock.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 295 AA;  
 Query Match 100.0%; Score 131; DB 16; Length 295;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-07;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGXPDEGKRGDACEGDSGSPFV 23  
 DB 224 AGXPDEGKRGDACEGDSGSPFV 246  
 RESULT 15  
 ID AAR74779 standard; Protein; 295 AA.  
 XX  
 AC AAR74779;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 04-NOV-1995 (first entry)  
 XX  
 DE Mutant thrombin E229S.  
 XX  
 KW Thrombin; oligonucleotide-directed mutagenesis; procoagulant;  
 KW anticoagulant; protein engineering; ss.  
 XX  
 CS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 265 /note= "Glu in wild-type"  
 FT Protein 37..295 /note= "mature protein"  
 XX  
 PN W09513385-A2.  
 XX  
 PD 18-MAY-1995.  
 XX  
 PF 14-NOV-1994; 94WO-US13104.  
 XX  
 PR 10-JUN-1994; 94US-0258038.  
 PR 12-NOV-1993; 93US-0152657.  
 XX  
 PA (GILE-) GILEAD SCI.

Result No.	Query Score	Match Length	DB	ID	Description
1	131	100.0	622	1	1BHU
2	127	96.9	236	2	C42696
3	124	94.7	625	1	TBBO
4	119	90.1	234	2	F42696
5	113	86.3	235	2	D42696
6	113	86.3	235	2	E42696
7	110	84.0	236	2	I42696
8	109	83.2	239	2	G42696
9	102	77.9	617	2	S10511
10	102	77.9	618	2	A35827
11	89	67.9	235	2	H42696
12	71.5	54.6	417	1	S00845
13	71	54.2	461	1	KHM
					protein C (activated)

14	70.5	53.8	482	1	EXRT	coagulation factor
15	70.5	53.8	638	1	KOHUP	plasma kallikrein
16	69.5	53.1	275	2	S40007	trypsin (EC 3.4.21
17	69.5	53.1	1524	2	T30337	polyprotein - Afri
18	68.5	52.3	161	2	162744	coagulation factor
19	68.5	52.3	488	1	EXHU	coagulation factor
20	68.5	52.3	1019	2	A38738	coagulation factor
21	67.5	51.5	161	2	148158	coagulation factor
22	67.5	51.5	282	2	184621	coagulation factor
23	67.5	51.5	459	2	J00419	coagulation factor
24	67.5	51.5	475	1	EXCH	coagulation factor
25	67.5	51.5	638	1	KOMSL	plasma kallikrein
26	67	51.1	225	2	S45356	proteolytic serine pr
27	67	51.1	264	2	S32794	trypsin-like prote
28	66.5	50.8	309	2	B49878	coagulation factor
29	66.5	50.8	1004	2	T30338	coagulation factor
30	66.5	50.0	267	2	S40006	trypsin (EC 3.4.21
31	65.5	50.0	274	2	S55339	trypsin (EC 3.4.21
32	65.5	50.0	275	2	S40005	trypsin (EC 3.4.21
33	65.5	50.0	277	2	S35340	trypsin (EC 3.4.21
34	65.5	50.0	638	1	KORTEL	plasma kallikrein
35	64.5	49.2	237	2	S55378	serine proteinase
36	64.5	49.2	238	1	TRW5Y	trypsin-like prote
37	64	48.9	191	2	S54115	complement factor
38	64	48.9	246	1	DBHU	complement factor
39	64	48.9	456	1	KXBO	protein C (activat
40	64	48.9	2616	2	A57096	nudel, protein prec
41	63.5	48.5	625	1	KFXU1	coagulation factor
42	63	48.1	461	1	JX0210	protein C (activat
43	62.5	47.7	375	1	A23689	limulus clotting e
44	62.5	47.7	416	1	S33777	hepsin (EC 3.4.21.
45	62.5	47.7	492	1	EXBO	coagulation factor

#### ALIGNMENTS

RESULT 1

THBU  
thrombin (EC 3.4.21.5) precursor [validated] - human  
N/Alternate names: coagulation factor II  
N/Contents: prothrombin  
C/Species: Homo sapiens (man)  
C/Date: 30-Nov-1980 #sequence, revision 22-Jul-1994 #text change 08-Dec-2000  
C/Accession: A29351; A00914; B00914; A37549; A37550; I51952  
R/Degen, S.J.F.; Davie, E.W., 1987  
Biochemistry 26, 6163-6177, 1987  
A/Title: Nucleotide sequence of the gene for human prothrombin.  
A/Reference number: A29351; MUID:8807877; PMID:2825773  
A/Accession: A29351  
A/Molecule type: DNA  
A/Residues: 1-622 <DE3>  
A/Cross-references: GB:M7262; GB:M33691; NID:g558069; PIND:AA63054.1;  
PIND:g339641  
R/Degen, S.J.F.; MacGillivray, R.T.A.; Davie, E.W.  
Biochemistry 22, 2087-2097, 1983

A/Title: Characterization of the complementary deoxyribonucleic acid and gene coding for human prothrombin.  
A/Reference number: A00914; MUID:83231469; PMID:6305407  
A/Accession: A00914  
A/Molecule type: mRNA  
A/Residues: 8-163, 'N', 165-622 <DE2>  
A/Cross-references: GB:V00595; GB:J00307; NID:g37128; PIND:CA22842.1;  
PIND:g1335344  
A/Accession: B00914  
A/Molecule type: DNA  
A/Residues: 188-311 <DE3>  
R/Maltz, D.A.; Hewett-Emmett, D.; Seeger, W.H.  
Proc. Natl. Acad. Sci. U.S.A. 74, 1969-1972, 1977  
A/Reference number: A37549; MUID:77193964; PMID:266717  
A/Accession: A37549  
A/Molecule type: protein  
A/Residues: 44-118, 'N', 120, 'S', 122-163, 'I', 165-175, 'A', 177-182, 'T', 184-193, 'M', 196-308, 'E', 309-314 <MAL>  
R/Butkowski, R.J.; Ellison, J.; Downing, M.R.; Mann, K.G.  
J. Biol. Chem. 252, 4942-4957, 1977  
A/Title: Primary structure of human prothrombin 2 and alpha-thrombin.  
A/Reference number: A37550; MUID:7707112; PMID:873923  
A/Accession: A37550  
A/Molecule type: protein  
A/Residues: 315-334, 'N', 336-348, 'N', 350-368, 'N', 370-397, 'N', 399-413, 'N', 415-484, 'N', 486-493, 'G', 495-503, 'V', 505-508, 'S', 510, 'V', 512-513, 'D', 515-528, 'AL', 531, 'Q', 533-622 <BUT>  
R/Rabier, M.; Blashill, A.; Furie, B.; Furie, B.C.  
J. Biol. Chem. 261, 13210-13215, 1986  
A/Reference number: A37551; MUID:87008532; PMID:3759958  
A/Contents: annotation; activation cleavages  
R/MacGillivray, R.T.; Irwin, D.M.; Guinto, E.R.; Stone, J.C.  
Ann. N. Y. Acad. Sci. 485, 73-79, 1986  
A/Title: Recombinant genetic approaches to functional mapping of thrombin.  
A/Reference number: I51952; MUID:87182874; PMID:3471151  
A/Accession: I51952  
A/Status: translated from GB/EMBL/DDBJ  
A/Molecule type: mRNA  
A/Residues: 1-2, 'R', 'S', 5-100 <RES>  
A/Cross-references: GB:N33031; NID:g190723; PIND:AAA60220.1; PID:g190724  
C/Comment: Thrombin, which cleaves bonds after Arg and Lys, converts fibrinogen to fibrin and activates factors V, VIII, XIII, and, in complex with thrombomodulin, protein C.  
C/Comment: Prothrombin is activated on the surface of a phospholipid membrane that binds the amino end of prothrombin and factors Va and Xa in calcium-dependent interactions. The activation peptide(s) can be removed either by factor Xa or thrombin; the cleavage into light and heavy chains is by factor Xa. It is not known whether one or two smaller activation peptides, with additional cleavage after 314-Arg, are released in natural blood clotting.  
C/Comment: The cleavage after Arg-198, observed in vitro, does not occur in plasma.  
C/Comment: The gamma-carboxyglutamic acid residues bind calcium ions, result from the carboxylation of glutamic acid residues by microsomal vitamin K-dependent carboxylase, and are necessary for calcium-dependent interaction with the negatively charged phospholipid membrane surface.  
C/Comment: The prothrombin precursor is synthesized in the liver.  
C/Genetics:

A/Gene: GDB:F2  
 A/Cross-references: GDB:119894; OMIM:176930  
 A/Map position: 11p11-11q12  
 A/Intons: 27/1; 80/3; 89/1; 106/1; 141/2; 187/1; 292/1; 335/1; 377/2; 433/2; 491/2; 552/1; 575/3  
 C/Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology  
 C/Keywords: acute phase; blood coagulation; calcium binding; carboxylglutamic acid; duplication; glycoprotein; hydrolase; kringle; liver; plasma; serine protease  
 F/1-24/Domain: signal sequence #status predicted <SIG>  
 F/25-43/Domain: propeptide #status predicted <PRO>  
 F/26-87/Domain: Gla domain homology <GLA>  
 F/44-622/Product: prothrombin #status experimental <PAT>  
 F/44-327/Domain: activation peptide #status experimental <APT>  
 F/108-186/Domain: kringle homology <KR1>  
 F/213-291/Domain: kringle homology <KR2>  
 F/328-363/Product: thrombin light chain #status experimental <LCR>  
 F/364-622/Product: thrombin heavy chain #status experimental <HCH>  
 F/364-613/Domain: trypsin homology <TRY>  
 F/49,50,57,59,62,63,66,69,72,75/Modified site: gamma-carboxylglutamic acid (Glu) #status experimental  
 F/60-65,90-103,108-186,129-169,157-181,213-291,234-274,262-286/Disulfide bonds: #status predicted  
 F/121,143/Binding site: carboxylate (Asn) (covalent) #status predicted  
 F/336-482,536-550,564-594/Disulfide bonds: #status predicted  
 F/391-407/Disulfide bonds: #status experimental  
 F/406,462/Active site: His, Asp #status predicted  
 F/416/Binding site: carboxylate (Asn) (covalent) #status experimental  
 F/568/Active site: Ser #status experimental  
  
 Query Match 100.0%; Score 131; DB 1; Length 622;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-10;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGKPEDEGRGDACEGDSGPFV 23  
 |||||  
 Db 551 AGKPEDEGRGDACEGDSGPFV 573

**RESULT 2**  
 C42696  
 Thrombin (EC 3.4.21.5) B chain - rabbit (fragment)  
 C/Species: Oryctolagus cuniculus (domestic rabbit)  
 C/Date: 26-May-1994 #sequence\_revision 26-May-1994 #text\_change 17-Mar-1999  
 C/Accession: C42696  
 R/Bentfield, D.K.; MacGillivray, R.T.A.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992  
 A/Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and sequence analysis of the B chain of thrombin from nine different species.  
 A/Reference number: A42696; XUID:92212913; PMID:1557383  
 A/Accession: C42696  
 A/Status: preliminary; nucleic acid sequence not shown; not compared with conceptual translation  
 A/Molecule type: mRNA  
 A/Residues: 1-236 <BAN>  
 A/Cross-references: GB:M81396  
 C/Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology

C/Keywords: hydrolase; serine proteinase  
 F/1-227/Domain: trypsin homology (fragment) <TRY>  
 Query Match 96.9%; Score 127; DB 2; Length 236;  
 Best Local Similarity 95.7%; Pred. No. 2.6e-10;  
 Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGKPEDEGRGDACEGDSGPFV 23  
 |||||  
 Db 165 AGKPEDEGRGDACEGDSGPFV 187

Search completed: February 11, 2004, 14:56:57  
 Job time : 16.5806 secs

OM protein - protein search, using sw model

Run on: February 11, 2004, 14:36:52 ; Search time 9.64516 Seconds

(without alignments)  
112.141 Million cell updates/sec

Title: US-10-05C-611-3

Perfect score: 131  
Sequence: 1 AGYKPEGRKRGACEDSGSPFV 23

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	131	100.0	622	1	THRB_HUMAN
2	124	94.7	625	1	THRB_BOVIN
3	102	77.9	617	1	THRB_RAT
4	102	77.9	618	1	THRB_MOUSE
5	73.5	56.1	290	1	MPN_HUMAN
6	71.5	54.6	417	1	HEPS_HUMAN
7	71.5	54.6	436	1	HEPS_MOUSE
8	71	54.2	161	1	PRTC_MACMU
9	71	54.2	461	1	PRTC_HUMAN
10	70.5	53.8	638	1	KAL_HUMAN
11	70	53.4	281	1	TRY2_DROER
12	69.5	53.1	275	1	TRY3_ANOGA
13	68.5	52.3	488	1	FA10_HUMAN
14	68.5	52.3	1019	1	LFC_CARRO
15	68.5	52.3	1019	1	LFC_TACRO
16	68	51.9	458	1	PRTC_RABIT
17	67.5	51.5	282	1	FA9_RAT

ALIGNMENTS

RESULT 1	THRB_HUMAN	STANDARD	PRTC	622 AA.
ID	THRB_HUMAN			
AC	P00734;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-JAN-1990 (Rel. 13, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Prothrombin precursor (EC 3.4.21.5) (Coagulation factor II).			
GN	F2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=88077877; PubMed=2825773;			
RA	Degen S.J.F., Davie E.W.;			
RT	"Nucleotide sequence of the gene for human prothrombin."			
RL	Nucleotide sequence of the gene for human prothrombin."			
RP	SEQUENCE FROM N.A. AND VARIANT MET-165.			
RA	Ridder M.J., Arnel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,			
RA	Ozuna M., Fosl C.L., Tsch E.J., Yi Q., Nickerson D.A.;			
RL	Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.			

18	67.5	51.5	459	1	FA9_MOUSE	P16294 mus musculu
19	67.5	51.5	475	1	FA10_CHICK	P23155 gallus gall
20	67.5	51.5	638	1	KAL_MOUSE	P26262 mus musculu
21	67	51.1	256	1	KLKE_HUMAN	Q92425 homo sapien
22	67	51.1	264	1	VDP_BOMMO	Q07943 bombyx mori
23	66.5	50.8	455	1	TMS5_MOUSE	Q9604 mus musculu
24	66.5	50.8	457	1	TMS5_HUMAN	Q9343 homo sapien
25	65.5	50.0	267	1	TRY7_ANOGA	P35041 anopheles g
26	65.5	50.0	274	1	TRY1_ANOGA	P35035 anopheles g
27	65.5	50.0	275	1	TRY4_ANOGA	P35038 anopheles g
28	65.5	50.0	277	1	TRY2_ANOGA	P35036 anopheles g
29	65.5	50.0	638	1	KAL_RAT	P14272 rattus norv
30	65	49.6	157	1	PRTC_CARPA	Q28278 carps famil
31	65	49.6	157	1	PRTC_CAPI	Q28315 capra hircu
32	65	49.6	157	1	PRTC_FELCA	Q28412 felis silve
33	65	49.6	157	1	PRTC_HORSE	Q28380 equus cabal
34	65	49.6	459	1	PRTC_PIG	Q93182 sus scrofa
35	64.5	49.2	238	1	TRY5_AEDAE	P29787 aedes aegyp
36	64.5	49.2	422	1	DEB1_HUMAN	Q9152 homo sapien
37	64.5	49.2	490	1	FA10_RABIT	O19045 oryctolagus
38	64	48.9	253	1	CPAD_HUMAN	P00746 homo sapien
39	64	48.9	259	1	CPAD_PIG	P51778 sus scrofa
40	64	48.9	456	1	PRTC_BOVIN	P00745 bos taurus
41	64	48.9	875	1	NEPR_HUMAN	P67330 homo sapien
42	64	48.9	2616	1	NDL_DROXE	P8159 drosophila
43	63.5	48.5	625	1	FA11_HUMAN	P03951 homo sapien
44	63	48.1	256	1	TRYE_DROER	P54627 drosophila
45	63	48.1	461	1	PRTC_MOUSE	P33587 mus musculu

RN [3]  
 RP SEQUENCE OF 8-622 FROM N.A.  
 RX MEDLINE=83231469; PubMed=6305407;  
 RA Degen S.J.F., McCallilivray R.T.A., Davie E.W.;  
 RT "Characterization of the complementary deoxyribonucleic acid and gene  
 RL coding for human prothrombin.";  
 RN Biochemistry 22:2087-2097(1983).  
 RN [4]  
 RP SEQUENCE OF 44-314.  
 RX MEDLINE=77193964; PubMed=266717;  
 RA Walz D.A., Hewett-Eames D., Seegers W.H.;  
 RT "Amino acid sequence of human prothrombin fragments 1 and 2.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 74:1969-1972(1977).  
 RN [5]  
 RP SEQUENCE OF 315-622.  
 RX MEDLINE=77207112; PubMed=873923;  
 RA Butkowski R.J., Elion J., Downing K.R., Mann K.G.;  
 RT "Primary structure of human prothrombin 2 and alpha-thrombin.";  
 RL J. Biol. Chem. 252:4942-4957(1977).  
 RN [6]  
 RP PROCESSING.  
 RX MEDLINE=87008332; PubMed=3759958;  
 RA Rabiet M.J., Blashill A., Furie B., Furie B.C.;  
 RT "Prothrombin fragment 1 X 2 X 3, a major product of prothrombin  
 RL activation in human plasma.";  
 RN J. Biol. Chem. 261:13210-13215(1986).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
 RX MEDLINE=90059942; PubMed=2583108;  
 RA Bode W., Mayr I., Baumann U., Huber R., Stone S.R., Hofsteenge J.;  
 RT "The refined 1.9 Å crystal structure of human alpha-thrombin:  
 RL interaction with D-Phe-Pro-Arg chloromethylketone and significance of  
 RT the Tyr-Pro-Pro-Tyr insertion segment.";  
 RL EXBO J. 8:3467-3475(1989).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE=90327074; PubMed=2374926;  
 RA Rydel T.J., Ravichandran K.G., Tulinsky A., Bode W., Huber R.,  
 RL Roltsch C., Fenton J.W. II;  
 RT "The structure of a complex of recombinant hirudin and human alpha-  
 RL thrombin.";  
 RN Science 249:277-280(1990).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
 RX MEDLINE=94350942; PubMed=8071320;  
 RA Rydel T.J., Yin W., Padmanabhan K.P., Blankenship D.T., Cardin A.D.,  
 RA Correa P.E., Fenton J.W. II, Tulinsky A.;  
 RT "Crystallographic structure of human gamma-thrombin.";  
 RL J. Biol. Chem. 269:22000-22006(1994).  
 RN [10]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE=97357286; PubMed=9214615;  
 RA van de Locht A., Bode W., Huber R., le Bonniec B.F., Stone S.R.,  
 RA Esmon C.T., Stubbs M.T.;  
 RT "The thrombin E192Q-BP1 complex reveals gross structural  
 RL rearrangements: implications for the interaction with antithrombin  
 RT and thrombomodulin.";  
 RN EXBO J. 16:2977-2984(1997).  
 RN [11]  
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 328-601.  
 RX MEDLINE=99162521; PubMed=10051558;  
 RA Quinto E.R., Caccia S., Rose T., Fueterer K., Wakaman G., di Cera E.;  
 RT "Unexpected crucial role of residue 225 in serine proteases.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:1852-1857(1999).  
 RN [12]  
 RP VARIANT BARCELONA.  
 RX MEDLINE=87033739; PubMed=3771562;  
 RA Rabiet M.-J., Furie B.C., Furie B.;  
 RT "Molecular defect of prothrombin Barcelona. Substitution of cysteine  
 RL for arginine at residue 273.";  
 RN J. Biol. Chem. 261:15045-15048(1986).  
 RN [13]  
 RP VARIANT FRANKFURT.  
 RX MEDLINE=59313001; PubMed=792730;  
 RA Degen S.J.F., McDowell S.A., Sparks L.M., Scharer I.;  
 RT "Prothrombin Frankfurt: a dysfunctional prothrombin characterized by  
 RL substitution of Glu-466 by Ala.";  
 RN Thromb. Haemost. 73:203-209(1995).  
 RN [14]  
 RP VARIANTS HIMI-1 AND HIMI-2.  
 RX MEDLINE=90043342; PubMed=1421398;  
 RA Morishita E., Saito M., Kumabeashi I., Asakura H., Matsuda T.,  
 RA Yamaguchi K.;  
 RT "Prothrombin Himi: a compound heterozygote for two dysfunctional  
 RL prothrombin molecules (Met-337-->Phe and Arg-388-->His).";  
 RN Blood 80:2275-2280(1992).  
 RN [15]  
 RP VARIANT PADUA-1.  
 RX MEDLINE=95169898; PubMed=7865694;  
 RA James H.L., Kim D.J., Zheng D.-Q., Girolami A.;  
 RT "Prothrombin Padua I: incomplete activation due to an amino acid  
 RL substitution at a factor Xa cleavage site.";  
 RN Blood Coagul. Fibrinolysis 5:841-844(1994).  
 RN [16]  
 RP VARIANT QUICK-1.  
 RX MEDLINE=89207504; PubMed=3242619;  
 RA Henriksen R.A., Mann K.G.;  
 RT "Identification of the primary structural defect in the dysprothrombin  
 RL thrombin Quick I: substitution of cysteine for arginine-382.";  
 RN Biochemistry 27:9160-9165(1988).  
 RN [17]  
 RP VARIANT QUICK-2.  
 RX MEDLINE=89247398; PubMed=2719946;  
 RA Henriksen R.A., Mann K.G.;  
 RT "Substitution of valine for glycine-559 in the congenital dysprothrombin  
 RL thrombin Quick II alters primary substrate specificity.";  
 RN Biochemistry 28:2078-2082(1989).  
 RN [18]  
 RP VARIANT SALAKTA.  
 RX MEDLINE=92378979; PubMed=1354985;  
 RA Miyata T., Ayuga R., Uneyama H., Beseaud A., Guillouin M.-C.,  
 RA Iwanaga S.;  
 RT "Prothrombin Salakta: substitution of glutamic acid-466 by alanine  
 RL reduces the fibrinogen clotting activity and the esterase activity.";  
 RN

RL Biochemistry 31:7457-7462(1992).  
 RN [19]  
 RP VARIANT TOKUSHIMA.  
 RX MEDLINE=87185407; PubMed=3567158;  
 RA Miyata T., Morita T., Inomoto T., Kawachi S., Shirakami A.,  
 RA Iwanaga S.;  
 RT "Prothrombin Tokushima, a replacement of arginine-418 by tryptophan  
 RT that impairs the fibrinogen clotting activity of derived thrombin  
 RT Tokushima.";  
 RL Biochemistry 26:1117-1122(1987).  
 RN [20]  
 RP VARIANT TOKUSHIMA.  
 RX MEDLINE=87101511; PubMed=3801671;  
 RA Inomoto T., Shirakami A., Kawachi S., Shigeakiyo T., Saito S.,  
 RA Miyoshi K., Morita T., Iwanaga S.;  
 RT "Prothrombin Tokushima: characterization of dysfunctional thrombin  
 RT derived from a variant of human prothrombin.";  
 RL Blood 69:565-569(1987).  
 RN [21]  
 RP VARIANT TOKUSHIMA.  
 RX MEDLINE=92256895; PubMed=1349838;  
 RA Iwahana H., Yoshimoto K., Shigeakiyo T., Shirakami A., Saito S.,  
 RA Itakura M.;  
 RT "Detection of a single base substitution of the gene for prothrombin  
 RT Tokushima. The application of PCR-SSCP for the genetic and molecular  
 RT analysis of dysprothrombinemia.";  
 RL Int. J. Hematol. 53:193-100(1992).  
 RN [22]  
 RP VARIANT TYPE-3.  
 RX MEDLINE=83204687; PubMed=6405779;  
 RA Board P.G., Shaw D.C.;  
 RT "Determination of the amino acid substitution in human prothrombin  
 RT type 3 (157 Glu leads to Lys) and the localization of a third  
 RT thrombin cleavage site.";  
 RL Br. J. Haematol. 54:245-254(1983).  
 RN [23]  
 RP VARIANTS ME1-165 AND THR-386.  
 RX MEDLINE=9318093; PubMed=10391209;  
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patti N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.O.,  
 RA Lander E.S.;  
 RT "Characterization of single-nucleotide polymorphisms in coding regions  
 RT of human genes.";  
 RL Nat. Genet. 22:231-238(1999).  
 RN [24]  
 RP ERRATUM.  
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patti N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.O.,  
 RA Lander E.S.;  
 RL Nat. Genet. 23:373-373(1999).  
 CC -1- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS  
 CC FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII,  
 CC AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.  
 CC -1- CATALYTIC ACTIVITY: preferential cleavage: Arg-|-Gly; activates  
 CC fibrinogen to fibrin and releases fibrinopeptide A and B.

CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN THE LIVER, FOUND IN PLASMA.  
 CC -1- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,  
 CC RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MITOCHONDRIAL  
 CC ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES  
 CC ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY  
 CC CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION  
 CC  
 Query Match 100.0%; Score 131; DB 1; Length 622;  
 Best local similarity 100.0%; Pred. No. 2,1e-10;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGKPEGRGDAGCGSGSPV 23  
 Db 551 AGKPEGRGDAGCGSGSPV 573  
 RESULT 2  
 ID THRB\_BOVIN STANDARD; PRT; 625 AA.  
 AC P00735;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Prothrombin precursor (EC 3.4.21.5).  
 GN F2.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OC NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88245190; PubMed=3379642;  
 RA Irwin D.M., Robertson K.A., Macgillivray R.T.A.;  
 RT "Structure and evolution of the bovine prothrombin gene.";  
 RL J. Mol. Biol. 200:31-45(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84203525; PubMed=6326805;  
 RA Macgillivray R.T.A., Davie E.W.;  
 RT "Characterization of bovine prothrombin mRNA and its translation  
 RT product.";  
 RL Biochemistry 23:1626-1634(1984).  
 RN [3]  
 RP SEQUENCE OF 44-625, DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.  
 RA Magnusson S., Sottrup-Jensen L., Petersen T.E., Claess H.;  
 RL (In) Hemker H.C., Veitkamp J.J. (eds.);  
 RL Biochimica Symposium on prothrombin and related coagulation factors,  
 RL pp.25-46, Leiden University Press, Leiden (1975).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF ACTIVATION PEPTIDE 1.  
 RX MEDLINE=86296631; PubMed=3741841;  
 RA Park C.H., Tulinsky A.;  
 RT "Three-dimensional structure of the kinkle sequence: structure of  
 RT prothrombin fragment 1.";  
 RL Biochemistry 25:3977-3982(1986).

RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF ACTIVATION PEPTIDE 1.  
 RX MEDLINE=91311686; PubMed=1956869;  
 RA Seshadri T.-P., Tulinsky A., Skrzypczak-Janekun E., Park C.H.;  
 RT "Structure of bovine prothrombin fragment 1 refined at 2.25-A  
 resolution.";  
 RL J. Mol. Biol. 220:481-494(1991).  
 RN [6]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF ACTIVATION PEPTIDE 1.  
 RX MEDLINE=92190185; PubMed=147238;  
 RA Soriano-garcia M., Padmanabhan K., de Vos A.M., Tulinsky A.;  
 RT "The Ca2+ ion and membrane binding structure of the Gla domain of Ca-  
 prothrombin fragment 1.";  
 RL Biochemistry 31:2554-2566(1992).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE=92218459; PubMed=1560020;  
 RA Martin P.D., Robertson W., Turk D., Huber R., Edwards B.F.P.;  
 RT "The structure of residues 7-16 of the A alpha-chain of human  
 fibrinogen bound to bovine thrombin at 2.3-A resolution.";  
 RL J. Biol. Chem. 267:7911-7920(1992).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE=92389319; PubMed=1518046;  
 RA Brandstetter H., Turk D., Hoeffken H.W., Grosse D., Stuerzebecher J.,  
 RA Martin P.D., Edwards B.F.P., Bode W.;  
 RT "Refined 2.3 A X-ray crystal structure of bovine thrombin complexes  
 formed with the benzamide and arginine-based thrombin inhibitors  
 NAPAP, 4-TAAP and MQPA. A starting point for improving  
 antithrombotics.";  
 RL J. Mol. Biol. 226:1085-1089(1992).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF COMPLEX WITH ORNITHODORIN.  
 RX MEDLINE=97102783; PubMed=8947023;  
 RA van de Locht A., Stubbs M.T., Bode W., Friedrich T., Bollschweiler C.,  
 RA Hoeffken W., Huber R.;  
 RT "The ornithodorin-thrombin crystal structure, a key to the TAP  
 enigma?"  
 RL EMBO J. 15:6011-6017(1996).  
 RN [10]  
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH TRIPAIN.  
 RX MEDLINE=98004486; PubMed=9342325;  
 RA Fuentes-Prior P., Noeske-Jungblut C., Donner P., Schleuning W.D.,  
 RA Huber R., Bode W.;  
 RT "Structure of the thrombin complex with triapin, a lipocalin-like  
 exosite-binding inhibitor derived from a tritacrine bug.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:11645-11650(1997).  
 RN [11]  
 RP GENE STRUCTURE.  
 RX MEDLINE=86077733; PubMed=3000440;  
 RA Irwin D.M., Ahern K.G., Pearson G.D., McGillivray R.T.A.;  
 RT "Characterization of the bovine prothrombin gene.";  
 RL Biochemistry 24:6854-6861(1985).  
 CC -1- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS  
 CC FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII,  
 CC AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.  
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Gly; activates

CC fibrinogen to fibrin and releases fibrinopeptide A and B.  
 CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN THE LIVER, FOUND IN PLASMA.  
 CC -1- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,  
 CC RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOMAL  
 CC ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES  
 CC ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY  
 CC CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION  
 CC OF PROTHROMBIN TO THROMBIN.  
 CC -1- MISCELLANEOUS: PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A  
 CC PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN &  
 CC FACTORS VA & XA IN CA-DEPENDENT INTERACTIONS; FACTOR XA REMOVES  
 CC THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT &  
 CC HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR  
 CC V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF  
 CC THROMBIN.  
 CC -1- MISCELLANEOUS: THROMBIN CAN ITSELF CLEAVE THE AMINO TERMINAL  
 CC FRAGMENT (FRAGMENT 1) OF THE PROTHROMBIN, PRIOR TO ITS ACTIVATION  
 CC BY FACTOR XA.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -1- DATABSE: NAME=Prozyme technical fact sheet;  
 CC WWW="http://www.prozyme.com/technical/thrombindata.html".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: V00135; CAA23451.1; -;  
 CC EMBL: J00041; AAA30761.1; -;  
 CC PIR: S02337; TBSO.  
 CC PDB: 1BBR; 31-JAN-94.  
 CC PDB: 1ETR; 31-JAN-94.  
 CC PDB: 1ETR; 31-JAN-94.  
 CC PDB: 1ERT; 31-JAN-94.  
 CC PDB: 1HRT; 31-JAN-94.  
 CC PDB: 1HRT; 31-JAN-94.  
 CC PDB: 2PT1; 31-JAN-94.  
 CC PDB: 2PT2; 31-JAN-94.  
 CC PDB: 2SPT; 31-MAY-94.  
 CC PDB: 1KRW; 07-JUL-97.  
 CC PDB: 1KRW; 07-JUL-97.  
 CC PDB: 1TBO; 14-OCT-96.  
 CC PDB: 1TBR; 14-OCT-96.  
 CC PDB: 1TBC; 23-JUL-97.  
 CC PDB: 1VTI; 21-APR-97.  
 CC PDB: 1YCP; 06-MAY-98.  
 CC PDB: 1A0H; 17-JUN-98.  
 CC PDB: 1AVG; 16-FEB-99.  
 CC PDB: 1ERT; 24-DEC-97.  
 CC PDB: 1IDS; 12-SEP-01.  
 CC PDB: 1UVT; 19-NOV-97.  
 CC PDB: 2HEP; 31-JAN-94.  
 CC MEMOPS: S01.217; -.

DR	InterPro: IPR001314; Chymotrypsin.
DR	InterPro: IPR002383; GLA_blood.
DR	InterPro: IPR000001; Kringle.
DR	InterPro: IPR003966; Prothrombin.
DR	InterPro: IPR001254; Ser_protease_Try.
DR	InterPro: IPR000294; VitK_dep_GLA.
DR	Pfam: PF00594; Gla; 1.
DR	Pfam: PF00051; kringle; 2.
DR	Pfam: PF00089; trypsin; 1.
DR	PRINTS: PR00722; CHYMOTRYPSIN.
DR	PRINTS: PR00001; GLABLOOD.
DR	PRINTS: PR00018; KRINGLE.
DR	PRINTS: PR01505; PROTHROMBIN.
DR	ProDom: PD000395; Kringle; 2.
DR	SMART: SM00069; GLA; 1.
DR	SMART: SM00130; KR; 2.
DR	SMART: SM00020; TRYP_SPE; 1.
DR	PROSITE: PS00011; GLU_CARBOXYLATION; 1.
DR	PROSITE: PS00021; KRINGLE_1; 2.
DR	PROSITE: PS00070; KRINGLE_2; 2.
DR	PROSITE: PS00240; TRYPSIN_DOM; 1.
DR	PROSITE: PS00134; TRYPSIN_HIS; 1.
DR	PROSITE: PS00135; TRYPSIN_SER; 1.
KW	Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;
KW	Vitamin K; Zymogen; Gamma-carboxyglutamic acid; Acute phase; Liver;
KW	Hydrolase; Serine protease; Kringle; Signal; 3D-structure.
FT	SIGNAL 1 24 POTENTIAL.
FT	PROPEP 25 43
FT	CHAIN 44 625 PROTHROMBIN.
FT	PEPTIDE 44 199 ACTIVATION PEPTIDE (FRAGMENT 1).
FT	PEPTIDE 200 317 ACTIVATION PEPTIDE (FRAGMENT 2).
FT	CHAIN 318 366 THROMBIN LIGHT CHAIN (A).
FT	CHAIN 367 625 THROMBIN HEAVY CHAIN (B).
FT	DOMAIN 109 187 KRINGLE 1.
FT	DOMAIN 214 292 KRINGLE 2.
FT	DOMAIN 367 625 SERINE PROTEASE.
FT	SITE 199 200 CLEAVAGE (BY THROMBIN).
FT	SITE 317 318 CLEAVAGE (BY FACTOR XA).
FT	SITE 366 367 CLEAVAGE (BY FACTOR XA).
FT	ACT_SITE 409 409 CHANGE RELAY SYSTEM.
FT	ACT_SITE 465 465 CHANGE RELAY SYSTEM.
FT	ACT_SITE 571 571 CHANGE RELAY SYSTEM.
FT	MOD_RES 50 50 GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES 51 51 GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES 58 58 GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES 60 60 GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES 63 63 GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES 64 64 GAMMA-CARBOXYGLUTAMIC ACID.
Query Match 94.7%; Score 124; DB 1; Length 625;	
Best Local Similarity 95.7%; Pred. No. 1.9e-09;	
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	1 AGYRDEGRGADACEGDSGGPFV 23
DB	
	554 AGYRDEGRGADACEGDSGGPFV 576

OM protein - protein search, using sw model

Run on: February 11, 2004, 14:47:57 ; Search time 39.3226 Seconds

(without alignments)  
150,936 Million cell updates/sec

Title: US-10-050-611-3

Perfect score: 131

Sequence: 1 AGYKDEGRKADCEGDSGSPFV 23

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 segs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seg length: 0

Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: SPRENEL 23:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mhcc:\*  
9: sp\_organelle:\*  
10: sp\_phage:\*  
11: sp\_plant:\*  
12: sp\_rodent:\*  
13: sp\_virus:\*  
14: sp\_vertebrate:\*  
15: sp\_unclassified:\*  
16: sp\_virus:\*  
17: sp\_bacteriap:\*  
18: sp\_bacteriap:\*  
19: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result Query  
No. Score Match Length DB ID Description

Search completed: February 11, 2004, 14:56:05  
Job time : 40.3226 secs

1	127	96.9	235	6	Q28731	Q28731 oryctolagus
2	118	90.1	235	13	Q90387	Q90387 cymops pyrr
3	113	86.3	235	13	Q91004	Q91004 gekko gekko
4	113	86.3	607	13	Q91001	Q91001 gallus gall
5	113	86.3	608	13	Q91007	Q91007 struthio ca
6	109	83.2	239	13	Q91218	Q91218 oncorhynch
7	105	80.2	420	13	Q90504	Q90504 eptatretus
8	98	74.8	172	13	Q90501	Q90501 eptatretus
9	92	70.2	234	13	Q90244	Q90244 actinenser t
10	72.5	55.3	389	13	Q9PVX7	Q9PVX7 xenopus lae
11	72.5	55.3	374	13	Q90WD8	Q90WD8 bufo japoni
12	71.5	54.6	435	11	Q9CM97	Q9CM97 mus masculu
13	71.5	54.6	799	11	Q9PD10	Q9PD10 mus masculu
14	71.5	54.6	802	4	Q8IU52	Q8IU52 homo sapien
15	71.5	54.6	811	4	Q8IU50	Q8IU50 homo sapien
16	71	54.2	195	4	Q8J008	Q8J008 homo sapien
17	71	54.2	195	4	Q8J007	Q8J007 homo sapien
18	71	54.2	195	4	Q8J006	Q8J006 homo sapien
19	71	54.2	195	4	Q8IX84	Q8IX84 homo sapien
20	71	54.2	211	4	Q8J009	Q8J009 homo sapien
21	70.5	53.8	161	11	Q63109	Q63109 rattus norv
22	70.5	53.8	259	5	Q9XV61	Q9XV61 ctanocephal
23	70.5	53.8	267	5	Q9BK47	Q9BK47 ludia foli
24	70.5	53.8	481	11	Q54740	Q54740 mus masculu
25	70.5	53.8	481	11	Q99L32	Q99L32 mus masculu
26	70.5	53.8	481	11	Q88B47	Q88B47 mus masculu
27	70.5	53.8	482	11	Q63207	Q63207 rattus norv
28	70	53.4	378	5	Q8SY50	Q8SY50 drosophila
29	69.5	53.1	200	11	Q92406	Q92406 mus masculu
30	69.5	53.1	1524	13	Q91674	Q91674 xenopus lae
31	68.5	52.3	161	6	Q28511	Q28511 macaca mula
32	68.5	52.3	236	5	Q9YV43	Q9YV43 schistosoma
33	68.5	52.3	488	5	Q9YV44	Q9YV44 schistosoma
34	68.5	52.3	766	4	Q8NBV4	Q8NBV4 homo sapien
35	68.5	52.3	1019	5	Q8T9S1	Q8T9S1 taeniopis
36	68.5	52.3	1083	5	Q26423	Q26423 carolinosc
37	68	51.9	686	13	Q9DGC2	Q9DGC2 cyrtinus ca
38	67.5	51.5	156	5	Q16007	Q16007 schistosoma
39	67.5	51.5	161	11	Q60546	Q60546 mesocricetu
40	67.5	51.5	264	5	Q02569	Q02569 culx quin
41	67.5	51.5	328	11	Q8BJR6	Q8BJR6 mus masculu
42	67.5	51.5	370	5	Q9VY44	Q9VY44 drosophila
43	67.5	51.5	387	5	Q9XV57	Q9XV57 ctanocephal
44	67.5	51.5	474	13	Q8JHC8	Q8JHC8 brachydanio
45	67.5	51.5	638	11	Q8R0P5	Q8R0P5 mus masculu

OM protein - protein search, using sw model

Run on: February 11, 2004, 14:35:52 ; Search time 49.7097 Seconds

(without alignments)  
73.441 Million cell updates/sec

Title: US-10-050-611-4

Perfect score: 131

Sequence: 1 AGYRDEGRKRDACEGDSGPRV 23

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107663 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107663

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

ALIGNMENTS

Result No.	Score	Match	Length	DB ID	Description
1	131	100.0	23	AAW63414	Cell growth/adhesi
2	131	100.0	23	AAH12893	Nerve tissue regen
3	131	100.0	23	AAH70363	Human thrombin rec
4	131	100.0	23	AAE2563	Human thrombin big
5	131	100.0	23	AAE20159	Human thrombin pep
6	131	100.0	23	AAH0898	Thrombin-derived p
7	131	100.0	116	AAW9115	Human zeta 2 preth
8	131	100.0	259	AAH11545	Human thrombin Asn
9	131	100.0	259	AAH60563	Human thrombin var
10	131	100.0	259	AAH60565	Human thrombin var
11	131	100.0	295	AAH74775	Wild-type thrombin
12	131	100.0	295	AAH74776	Mutant thrombin K5
13	131	100.0	295	AAH74777	Mutant thrombin E2
14	131	100.0	295	AAH74778	Mutant thrombin E2
15	131	100.0	295	AAH74779	Mutant thrombin E2
16	131	100.0	295	AAH74780	Mutant thrombin E2
17	131	100.0	295	AAH76033	Mutant thrombin E2
18	131	100.0	295	AAH76034	Mutant thrombin R2
19	131	100.0	295	AAH76035	Mutant thrombin R2
20	131	100.0	295	AAH76036	Mutant thrombin R2
21	131	100.0	295	AAH76037	Mutant thrombin W5
22	131	100.0	295	AAH76038	Mutant thrombin K5
23	131	100.0	295	AAH76039	Mutant thrombin W5
24	131	100.0	295	AAH76040	Mutant thrombin W5
25	131	100.0	295	AAW22892	Human mature throm
26	131	100.0	295	AAH08633	Amino acid sequenc
27	131	100.0	295	AAH60562	Human thrombin var
28	131	100.0	295	AAH60564	Human thrombin var
29	131	100.0	308	AAW99109	Human prothrombin
30	131	100.0	376	AAH41787	CD4/Thrombin fusio
31	131	100.0	376	AAH42789	Human CD4/thrombin
32	131	100.0	376	AAH10703	Human CD4-thrombin
33	131	100.0	579	AAH35763	Prothrombin (PT) .
34	131	100.0	579	AAH11546	Human prothrombin
35	131	100.0	579	AAH11544	Human prothrombin
36	131	100.0	579	AAW99108	Human prothrombin
37	131	100.0	615	AAH38741	Human prothrombin
38	131	100.0	615	AAH36216	Human prothrombin
39	131	100.0	615	AAH30377	Human prothrombin
40	131	100.0	622	AAH11543	Human prothrombin
41	131	100.0	622	AAH19566	Human prothrombin
42	131	100.0	622	AAH74671	Platelet membrane
43	131	94.7	111	AAW99113	Human F2 protein.
44	124	94.7	308	AAW99107	Bovine zeta 2 prot
45	124	94.7	582	AAW99106	Bovine prothrombin







Query Match 100.0%; Score 131; DB 23; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 3, 4e-08;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGYKPEDGKRGDACEGDSGSPFV 23  
 ||||||||||||||||||  
 DB 1 AGYKPEDGKRGDACEGDSGSPFV 23

RESULT 7  
 ID AAW99115 standard; protein; 116 AA.  
 XX  
 AC AAW99115;  
 XX  
 DT 14-MAY-1999 (first entry)  
 XX  
 DE Human zeta 2 prethrombin 2.  
 XX  
 KW Prothrombin; exosite assay; anticoagulant; blood clot; thrombin;  
 XX cardiovascular disease; stroke; haematological disorder.  
 OS Homo sapiens.  
 XX  
 XX W09855130-A1.  
 XX  
 PD 10-DEC-1998.  
 XX  
 PF 28-MAY-1998; 98WO-US10840.  
 XX  
 PR 08-APR-1998; 98US-0081030.  
 PR 06-JUN-1997; 97US-0048864.  
 XX  
 PA (UYEM-) UNIV EXORV.  
 XX  
 PI Krishnaswamy S;  
 XX  
 DR WPI, 1999-070237/06.  
 XX  
 PT Exosite assay for agents that inhibit catalytic cleavage of  
 PT prothrombin - at sites away from the active site of prothrombinase,  
 PT also new inhibitors, potentially useful as anticoagulants  
 XX  
 PS Disclosure; Page 44-45; 61pp; English.

CC An exosite assay has been developed for inhibition of the catalytic  
 CC cleavage of prothrombin (PTn) to thrombin (Th) by prothrombinase (I), at  
 CC a site remote from the catalytic site of (I) comprises: (a) preparing a  
 CC solution containing 0.05-20  $\mu$ M substrate (S), that includes a protease  
 CC cleavage site and exosite-binding determinant; 0.05-200 nM factor Va;  
 CC 30-500 micro M phospholipids (PL); test inhibitor (A) in buffer of pH  
 CC 7-9, containing 1-10 mM calcium ions but no calcium-chelating agent;  
 CC (b) initiating catalytic cleavage of (S) by adding an aliquot of factor  
 CC Va (to final concentration 0.05-200 nM) so that there is an excess of Va  
 CC over Xa, forming a S/(I) complex; (c) withdrawing aliquots of the  
 CC reaction mixture, quenching them; and (d) assaying for concentration of

CC Th. Alternatively, in the initial solution S is replaced by the same  
 CC concentration of Xa (less than the amount of Va), and reaction is started  
 CC by adding S. Also described in the present invention are inhibitors (A')  
 CC having IC50 less than 1  $\mu$ M identified by this assay. (A') are  
 CC potentially useful as a new class of anticoagulants for treatment of  
 CC cardiovascular disease, stroke and haematological disorders. The method  
 CC is based on the finding that exosite interactions are critical for  
 CC substrate specificity in catalytic formation of Th. The present sequence  
 CC represents human zeta 2 prethrombin 2.

SQ Sequence 116 AA;  
 XX

Query Match 100.0%; Score 131; DB 20; Length 116;  
 Best Local Similarity 100.0%; Pred. No. 1, 4e-07;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGYKPEDGKRGDACEGDSGSPFV 23  
 ||||||||||||||||||  
 DB 45 AGYKPEDGKRGDACEGDSGSPFV 67

RESULT 8  
 ID AAM11545 standard; Protein; 259 AA.  
 XX  
 AC AAM11545;  
 XX  
 DT 01-OCT-1997 (first entry)  
 XX  
 DE Human thrombin Asn99 mutant.  
 XX  
 KW Prothrombin; mutant; mutetin; platelet aggregation; blood clotting;  
 KW coagulation; reduce; decrease; hirudin; heparin; anti-thrombin III;  
 KW antagonist; D99N.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT 1..259 /label= thrombin\_Asn99  
 FT Protein  
 FT Misc-difference 99 /note= "Wild-type Asp residue has been replaced by  
 FT Asn"  
 FT  
 FN W09641868-A2.  
 XX  
 PD 27-DEC-1996.  
 XX  
 PF 12-JUN-1996; 96MO-AT00105.  
 XX  
 PR 13-JUN-1995; 95AT-0001006.  
 XX  
 PA (IMMO ) IMMUNO AG.  
 XX  
 PI Eidl J, Falkner F, Fischer B, Mitterer A, Schlokat U;

DR WPI; 1997-063455/06.  
 XX  
 CC Prothrombin mutants with reduced clotting activity - useful as  
 PT antagonists of thrombin inhibitors or for anticoagulant therapy  
 XX  
 PS Example 3; Page -; 73pp; German.  
 XX  
 CC Prothrombin mutants having one or more changes in amino acid sequence  
 CC compared with the natural protein and having 0-10% (preferably 0-0.25%)  
 CC of the activity of the natural protein are claimed, provided that the  
 CC changes in amino acid sequence do not affect the capacity of the  
 CC mutants to bind to specific ligands and receptors. The mutants have  
 CC greatly reduced clotting activity and are useful as antagonists of  
 CC thrombin inhibitors such as hirudin, heparin and anti-thrombin III.  
 CC The mutations may also result in changes to the in vivo half-life  
 CC of prothrombin. The half-life may be reduced to less than 10 minutes  
 CC or the mutant prothrombin may have an extended half-life of more than  
 CC 1 hour, making it useful as an anticoagulant and to inhibit side-  
 CC effects of anti-coagulant treatment. They are converted to inactive  
 CC thrombin and are able to compete with native, active thrombin for  
 CC binding to receptors. The present sequence represents the thrombin  
 CC mutant which is derived by trypsin cleavage of a specifically  
 CC claimed human prothrombin mutant in which Asp at position 419 is  
 CC changed to Asn. The thrombin Asn99 mutant was found to have only  
 CC 0.24% of the activity of wild-type thrombin on a chromogenic  
 CC substrate.  
 CC (Note: This sequence does not appear in the specification and has  
 CC been produced by modifying the wild-type sequence of human  
 CC prothrombin which appears in figure 1).  
 XX  
 SQ Sequence 259 AA:  
 QY 1 AGKRPDEGKRGDACEGDSGSPFV 23  
 DB 188 AGKRPDEGKRGDACEGDSGSPFV 210  
 ||||||||||||||||||||  
 RESULT 9  
 ABBP60563  
 ID ABBP60563 standard; protein; 259 AA.  
 XX  
 AC ABBP60563;  
 XX  
 DT 28-MAR-2003 (first entry)  
 XX  
 DE Human thrombin variant W215A B-chain.  
 XX  
 KW Human; thrombin; W215A; anticoagulant; prothrombin; antithrombotic;  
 XX thrombus; protein C activation.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers

FT Misc-difference 229 /note="Wild-type Trp substituted by Ala"  
 FT  
 XX  
 PN WO2002100337-A2.  
 XX  
 PD 19-DEC-2002.  
 XX  
 PF 07-JUN-2002; 2002WO-US18211.  
 XX  
 PR 08-JUN-2001; 2001US-297089P.  
 XX  
 PA (UYEK-) UNIV EMORY.  
 XX  
 PI Gruber A, Hanson SR, Di Cera E;  
 XX  
 DR WPI; 2003-156907/15.  
 XX  
 PT New variant thrombin, useful as an antithrombotic agent for inhibiting  
 PT the formation of a thrombus, for determining the level of protein C  
 PT activation in a blood sample, or for determining the thrombogenic  
 PT potential of a patient -  
 XX  
 PS Claim 15; Fig 2; 95pp; English.  
 XX  
 CC The invention relates to a novel variant human thrombin. The thrombin  
 CC variant of the invention has anticoagulant activity. The variant thrombin  
 CC or prothrombin is useful as an antithrombotic agent for inhibiting the  
 CC formation of a thrombus. The variant thrombin is also useful for  
 CC determining the level of protein C activation in a blood sample or the  
 CC thrombogenic potential of a patient. The present sequence represents the  
 CC B-chain of the thrombin variant W215A.  
 XX  
 SQ Sequence 259 AA:  
 QY 1 AGKRPDEGKRGDACEGDSGSPFV 23  
 DB 188 AGKRPDEGKRGDACEGDSGSPFV 210  
 ||||||||||||||||||||  
 RESULT 10  
 ABBP60565  
 ID ABBP60565 standard; protein; 259 AA.  
 XX  
 AC ABBP60565;  
 XX  
 DT 28-MAR-2003 (first entry)  
 XX  
 DE Human thrombin variant W215A/E217A B-chain.  
 XX  
 KW Human; thrombin; W215A/E217A; anticoagulant; prothrombin; antithrombotic;  
 XX thrombus; protein C activation.  
 OS Homo sapiens.

XX Key Location/Qualifiers  
FH Thrombin; oligonucleotide-directed mutagenesis; procoagulant;  
FT Misc-difference 227 /note= "Wild-type Trp substituted by Ala"  
FT FT Misc-difference 228 /note= "Wild-type Glu substituted by Ala"  
XX  
XX WO2002100337-A2.  
XX  
XX 19-DEC-2002.  
XX  
XX 07-JUN-2002; 2002NO-US18211.  
XX  
XX 08-JUN-2001; 2001US-297089P.  
XX  
XX (UYEM-) UNITV EMORY.  
XX  
XX Gruber A, Hanson SR, Di Cera E,  
PI WPI; 2003-156907/15.  
DR N-PSDB; AB225353.  
XX  
XX New variant thrombin, useful as an antithrombotic agent for inhibiting  
PT the formation of a thrombus, for determining the level of protein C  
PT activation in a blood sample, or for determining the thrombogenic  
PT potential of a patient -  
XX  
XX Claim 2; Fig 4; 95pp; English.  
PS  
XX The invention relates to a novel variant human thrombin. The thrombin  
CC variant of the invention has anticoagulant activity. The variant thrombin  
CC or prothrombin is useful as an antithrombotic agent for inhibiting the  
CC formation of a thrombus. The variant thrombin is also useful for  
CC determining the level of protein C activation in a blood sample or the  
CC thrombogenic potential of a patient. The present sequence represents the  
CC B-chain of the thrombin variant W215A/E217A (WE).  
XX  
SQ Sequence 259 AA;  
Query Match 100.0%; Score 131; DB 24; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.9e-07;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGKPDGKRGDACEGDSGSPFV 23  
DB 188 AGKPDGKRGDACEGDSGSPFV 210  
RESULT 11  
AAAT4775  
ID AAAT4775 standard; Protein; 295 AA.  
XX  
XX AAAT4775;  
AC  
XX 25-MAR-2003 (updated)  
DT 04-NOV-1995 (first entry)  
XX

DE Wild-type thrombin.  
XX  
XX Thrombin; oligonucleotide-directed mutagenesis; procoagulant;  
KW anticoagulant; protein engineering; sa.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Protein 37..295  
FT /note= "ature protein"  
XX  
XX WO9513385-A2.  
XX  
XX 18-MAY-1995.  
XX  
XX 14-NOV-1994; 94WO-US13104.  
XX  
XX 10-JUN-1994; 94US-0258038.  
PR 12-NOV-1993; 93US-0152657.  
XX  
XX (GILE-) GILEAD SCT.  
XX  
XX Gibbs CS, Leung LK, Tsaiang M;  
PI WPI; 1995-194103/25.  
DR N-PSDB; AAQ92455.  
XX  
XX Thrombin derive with segregated pro- and anticoagulant activities -  
PT useful for treating thrombotic disorders but also diagnosis,  
PT treatment of tumours, etc.  
XX  
XX  
XX Disclosure; Fig 1; 78pp; English.  
PS  
XX The sequence represents wild-type (reference) thrombin. Mutants  
CC of this sequence (AAAT4776-80 and AAAT6033-41) have at least 80%  
CC homology with thrombin, and are capable of protein-C activation  
CC without significant fibrinogen clotting activity, and vice versa  
CC (specifically have a ratio of protein-C activity to fibrinogen  
CC clotting activity of less than 0.5 or greater than 2 compared to  
CC thrombin). The mutant thrombin sequences, produced in recombinant  
CC cell culture or by in vitro methods, and are used to treat  
CC thrombotic conditions, particularly during cardiac bypass surgery  
CC and in cases of septic shock.  
XX  
XX (Updated on 25-MAR-2003 to correct PN field.)  
SQ Sequence 295 AA;  
Query Match 100.0%; Score 131; DB 16; Length 295;  
Best Local Similarity 100.0%; Pred. No. 3.3e-07;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGKPDGKRGDACEGDSGSPFV 23  
DB 224 AGKPDGKRGDACEGDSGSPFV 246  
RESULT 12

AA874776  
ID AA874776 standard; Protein; 295 AA.  
XX  
AC AA874776;  
XX  
DT 25-MAR-2003 (updated)  
DT 04-NOV-1995 (first entry)  
XX  
DE Mutant thrombin K52A, R233A.  
XX  
KW Thrombin; oligonucleotide-directed mutagenesis; procoagulant;  
KW anticoagulant; protein engineering; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 88 /note= "Lys in wild-type"  
FT Misc-difference 269 /note= "Arg in wild-type"  
FT Protein 37..295 /note= "mature protein"  
FT  
XX  
PN W09513385-A2.  
XX  
PD 18-MAY-1995.  
XX  
PF 14-NOV-1994; 94WO-US13104.  
XX  
PR 10-JUN-1994; 94US-0258038.  
PR 12-NOV-1993; 93US-0152657.  
XX  
PA (GILE-) GILEAD SCI.  
XX  
PI Gibbs CS, Leung LK, Tsang M;  
XX  
DR WPI; 1995-194103/25.  
XX  
PT Thrombin deriva with segregated pro- and anticoagulant activities -  
PT useful for treating thrombotic disorders but also diagnosis,  
PT treatment of tumours, etc.  
XX  
PS Claim 22; Page 63/3; 76pp; English.  
XX  
CC The mutant thrombin sequence, generated by oligonucleotide-directed  
CC mutagenesis, has at least 80% homology with thrombin, and is  
CC capable of protein-C activation without significant fibrinogen  
CC clotting activity, and vice versa (specifically, it has a ratio  
CC of protein-C activity to fibrinogen clotting activity of less than  
CC 0.5 or greater than 2 compared to thrombin). The mutant thrombin  
CC is produced in recombinant cell culture or by in vitro methods,  
CC and is used to treat thrombotic conditions, particularly during  
CC cardiac bypass surgery and in cases of septic shock.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 295 AA;

Query Match 100.0%; Score 131; DB 16; Length 295;  
Best Local Similarity 100.0%; Pred. No. 3..3e-07;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGYKPEGRGACGEGSGPFV 23  
DB 224 AGYKPEGRGACGEGSGPFV 246  
RESULT 13  
AA874777  
ID AA874777 standard; Protein; 295 AA.  
XX  
AC AA874777;  
XX  
DT 25-MAR-2003 (updated)  
DT 04-NOV-1995 (first entry)  
XX  
DE Mutant thrombin E229D.  
XX  
KW Thrombin; oligonucleotide-directed mutagenesis; procoagulant;  
KW anticoagulant; protein engineering; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 265 /note= "Glu in wild-type"  
FT Protein 37..295 /note= "mature protein"  
FT  
XX  
PN W09513385-A2.  
XX  
PD 18-MAY-1995.  
XX  
PF 14-NOV-1994; 94WO-US13104.  
XX  
PR 10-JUN-1994; 94US-0258038.  
PR 12-NOV-1993; 93US-0152657.  
XX  
PA (GILE-) GILEAD SCI.  
XX  
PI Gibbs CS, Leung LK, Tsang M;  
XX  
DR WPI; 1995-194103/25.  
XX  
PT Thrombin deriva with segregated pro- and anticoagulant activities -  
PT useful for treating thrombotic disorders but also diagnosis,  
PT treatment of tumours, etc.  
XX  
PS Claim 22; Page 63/3; 76pp; English.  
XX  
CC The mutant thrombin sequence, generated by oligonucleotide-directed  
CC mutagenesis, has at least 80% homology with thrombin, and is  
CC capable of protein-C activation without significant fibrinogen  
CC clotting activity, and vice versa (specifically, it has a ratio  
CC of protein-C activity to fibrinogen clotting activity of less than

CC 0.5 or greater than 2 compared to thrombin). The mutant thrombin  
 CC is produced in recombinant cell culture or by in vitro methods,  
 CC and is used to treat thrombotic conditions, particularly during  
 CC cardiac bypass surgery and in cases of septic shock.  
 CC (Updated on 25-MAR-2003 to correct FN field.)  
 XX  
 SQ Sequence 295 AA;  
 QY Query Match 100.0%; Score 131; DB 16; Length 295;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-07;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 DB 1 AGYKPDGKRGDACEGDSGGPFV 23  
 224 AGYKPDGKRGDACEGDSGGPFV 246  
 RESULT 14  
 ID AAR74778 standard; Protein; 295 AA.  
 AC AAR74778;  
 DT 25-MAR-2003 (updated)  
 DT 04-NOV-1995 (first entry)  
 DE Mutant thrombin E229F.  
 XX  
 KW Thrombin; oligonucleotide-directed mutagenesis; procoagulant;  
 KW anticoagulant; protein engineering; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 265 /note= "Glu in wild-type"  
 FT Protein 37..295 /note= "mature protein"  
 FT  
 XX  
 PN WO9513385-A2.  
 XX  
 PD 18-MAY-1995.  
 XX  
 PF 14-NOV-1994; 94WO-US13104.  
 XX  
 PR 10-JUN-1994; 94US-0258038.  
 PR 12-NOV-1993; 93US-0152657.  
 XX  
 PA (GILE-) GILEAD SCI.  
 XX  
 PI Gibbs CS, Leung LK, Tsiang M;  
 DR WPI; 1995-194103/25.  
 XX  
 PT Thrombin derive with segregated pro- and anticoagulant activities -  
 PT useful for treating thrombotic disorders but also diagnosis,  
 PT treatment of tumours, etc.

XX  
 PS Claim 22; Page 63/3; 78pp; English.  
 XX  
 CC The mutant thrombin sequence, generated by oligonucleotide-directed  
 CC mutagenesis, has at least 80% homology with thrombin, and is  
 CC capable of protein-C activation without significant fibrinogen  
 CC clotting activity, and vice versa (specifically, it has a ratio  
 CC of protein-C activity to fibrinogen clotting activity of less than  
 CC 0.5 or greater than 2 compared to thrombin). The mutant thrombin  
 CC is produced in recombinant cell culture or by in vitro methods,  
 CC and is used to treat thrombotic conditions, particularly during  
 CC cardiac bypass surgery and in cases of septic shock.  
 CC (Updated on 25-MAR-2003 to correct FN field.)  
 XX  
 SQ Sequence 295 AA;  
 QY Query Match 100.0%; Score 131; DB 16; Length 295;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-07;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 DB 1 AGYKPDGKRGDACEGDSGGPFV 23  
 224 AGYKPDGKRGDACEGDSGGPFV 246  
 RESULT 15  
 ID AAR74779 standard; Protein; 295 AA.  
 AC AAR74779;  
 DT 25-MAR-2003 (updated)  
 DT 04-NOV-1995 (first entry)  
 DE Mutant thrombin E229S.  
 XX  
 KW Thrombin; oligonucleotide-directed mutagenesis; procoagulant;  
 KW anticoagulant; protein engineering; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 265 /note= "Glu in wild-type"  
 FT Protein 37..295 /note= "mature protein"  
 FT  
 XX  
 PN WO9513385-A2.  
 XX  
 PD 18-MAY-1995.  
 XX  
 PF 14-NOV-1994; 94WO-US13104.  
 XX  
 PR 10-JUN-1994; 94US-0258038.  
 PR 12-NOV-1993; 93US-0152657.  
 XX  
 PA (GILE-) GILEAD SCI.

XX Gibbs CS, Leung LK, Tsiang M;  
PI WPI: 1995-194103/25.  
XX  
XX Thrombin derivs with segregated pro- and anticoagulant activities -  
XX useful for treating thrombotic disorders but also diagnosis,  
PT treatment of tumours, etc.  
XX  
XX Claim 22; Page 63/3; 78pp; English.  
XX  
XX The mutant thrombin sequence, generated by oligonucleotide-directed  
CC mutagenesis, has at least 80% homology with thrombin, and is  
CC capable of protein-C activation without significant fibrinogen  
CC clotting activity, and vice versa (specifically, it has a ratio  
CC of protein-C activity to fibrinogen clotting activity of less than  
CC 0.5 or greater than 2 compared to thrombin). The mutant thrombin  
CC is produced in recombinant cell culture or by in vitro methods,  
CC and is used to treat thrombotic conditions, particularly during  
CC cardiac bypass surgery and in cases of septic shock.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 295 aa;  
Query Match 100.0%; Score 131; DB 16; Length 295;  
Best Local Similarity 100.0%; Pred. No. 3.3e-07;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGKRPDEGKRGDACEGDSGSPFV 23  
|||||  
DB 224 AGKRPDEGKRGDACEGDSGSPFV 246

Search completed: February 11, 2004, 14:53:25  
Job time : 49.7097 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model  
Run on: February 11, 2004, 14:49:07 ; Search time 15.5806 Seconds  
(without alignments)  
141.963 Million cell updates/sec

Title: US-10-050-611-4  
Perfect score: 131  
Sequence: 1 AGKRPDEGKRGDACEGDSGSPFV 23

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: PIR:6:\*  
2: PIR1:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	131	100.0	622	1 TBRU	thrombin (EC 3.4.2
2	127	96.9	236	2 C42696	thrombin (EC 3.4.2
3	124	94.7	625	1 TBRU	thrombin (EC 3.4.2
4	118	90.1	234	2 D42696	thrombin (EC 3.4.2
5	113	86.3	235	2 D42696	thrombin (EC 3.4.2
6	113	86.3	235	2 E42696	thrombin (EC 3.4.2
7	110	84.0	236	2 I42696	thrombin (EC 3.4.2
8	109	83.2	239	2 G42696	thrombin (EC 3.4.2
9	102	77.9	617	2 S10511	thrombin (EC 3.4.2
10	102	77.9	618	2 A35827	thrombin (EC 3.4.2
11	89	67.9	235	2 H42696	thrombin (EC 3.4.2
12	71.5	54.6	417	1 S00845	hepsin (EC 3.4.21.
13	71	54.2	461	1 KXHO	protein C (activat

14	70.5	53.6	482	1	EXRT	coagulation factor
15	70.5	53.6	638	1	KOHLP	plasma kallikrein
16	69.5	53.1	275	2	S40007	trypsin (EC 3.4.21
17	69.5	53.1	1324	2	T30337	polyprotein - Afri
18	68.5	52.3	161	2	162744	coagulation factor
19	68.5	52.3	488	1	EXHU	coagulation factor
20	68.5	52.3	1019	2	A38738	coagulation factor
21	67.5	51.5	161	2	148158	coagulation factor
22	67.5	51.5	282	2	184621	coagulation factor
23	67.5	51.5	459	2	J00419	coagulation factor
24	67.5	51.5	475	1	EXCH	coagulation factor
25	67.5	51.5	638	1	KOWSPL	plasma kallikrein
26	67.5	51.5	225	2	S45356	protease, serine pr
27	67	51.1	264	2	S32794	trypsin-like prote
28	66.5	50.8	309	2	B49878	coagulation factor
29	66.5	50.8	1004	2	T30338	oviductin (EC 3.4.
30	65.5	50.0	287	2	S40006	trypsin (EC 3.4.21
31	65.5	50.0	274	2	S35339	trypsin (EC 3.4.21
32	65.5	50.0	275	2	S40005	trypsin (EC 3.4.21
33	65.5	50.0	277	2	S35340	trypsin (EC 3.4.21
34	65.5	50.0	638	1	KORPL	plasma kallikrein
35	64.5	49.2	237	2	S53378	serine proteinase
36	64.5	49.2	238	1	TRWV5Y	trypsin-like prote
37	64	48.5	191	2	S54115	complement factor
38	64	48.5	246	1	DMHU	complement factor
39	64	48.5	456	1	KRBO	protein C (activat
40	64	48.9	2616	2	A57096	nudel protein prec
41	63.5	48.5	625	1	KFHU1	coagulation factor
42	63	48.1	461	1	JX0210	protein C (activat
43	62.5	47.7	375	1	A23689	limulus clotting e
44	62.5	47.7	416	1	S33777	heparin (EC 3.4.21.
45	62.5	47.7	492	1	EXBO	coagulation factor

# ALIGNMENTS

**RESULT 1**  
**THBU**  
 thrombin (EC 3.4.21.5) Precursor [validated] - human  
 N/Alternate names: coagulation factor II  
 N/Contains: prothrombin  
 C/Species: Homo sapiens (man)  
 C/Date: 30-Nov-1980 #sequence revision 22-Jul-1994 #text change 08-Dec-2000  
 C/Accession: A29351, A00914, E00914, A37549, A37550, I51952  
 R/Degen, S.J.E.; Davie, E.W.  
 Biochemistry 26, 6165-6177, 1987  
 A/Title: Nucleotide sequence of the gene for human prothrombin.  
 A/Reference number: A29351; M0ID:8607877; PMID:2825773  
 A/Accession: A29351  
 A/Molecule type: DNA  
 A/Residues: 1-622 <DE3>  
 A/Cross-references: GB:M17262; GB:M33691; NID:9558069; PIND:AA63054.1;  
 PID:9339641  
 R/Degen, S.J.F.; MacGillivray, R.T.A.; Davie, E.W.  
 Biochemistry 22, 2087-2097, 1983

A/Title: Characterization of the complementary deoxyribonucleic acid and gene coding for human prothrombin.  
 A/Reference number: A00914; M0ID:83231469; PMID:6305407  
 A/Accession: A00914  
 A/Molecule type: mRNA  
 A/Residues: 6-163, 'N', 165-622 <DE2>  
 A/Cross-references: GB:V00595; GB:J00307; NID:937128; PIND:CAA23842.1; PID:9133544  
 A/Accession: E00914  
 A/Molecule type: DNA  
 A/Residues: 189-311 <DE3>  
 R/Walz, D.A.; Hewett-Emmett, D.; Seegers, W.H.  
 Proc. Natl. Acad. Sci. U.S.A. 74, 1968-1972, 1977  
 A/Reference number: A37549; M0ID:77193964; PMID:266717  
 A/Accession: A37549  
 A/Molecule type: Protein  
 A/Residues: 44-118, 'N', 120, 'S', 122-163, 'I', 165-175, 'A', 177-182, 'T', 184-193, 'M', 196-308, 'E', 309-314 <WAL>  
 R/Burkowski, R.J.; Ellison, J.; Downing, M.R.; Mann, K.G.  
 J. Biol. Chem. 252, 4942-4957, 1977  
 A/Title: Primary structure of human prothrombin 2 and alpha-thrombin.  
 A/Reference number: A37550; M0ID:7707112; PMID:873923  
 A/Accession: A37550  
 A/Molecule type: Protein  
 A/Residues: 315-334, 'N', 336-348, 'N', 350-368, 'N', 370-397, 'N', 399-413, 'N', 415-464, 'N', 466-493, 'G', 495-503, 'V', 505-508, 'S', 510, 'V', 512-513, 'D', 515-528, 'AL', 531, 'Q', 533-622 <BUR>  
 R/Rablier, M.C.; Blachin, A.; Furie, B.; Furie, B.C.  
 J. Biol. Chem. 261, 13210-13215, 1986  
 A/Reference number: A37551; M0ID:87008532; PMID:3759958  
 A/Contents: annotation; activation cleavages  
 R/MacGillivray, R.T.; Irwin, D.M.; Guinto, E.R.; Stone, J.C.  
 Ann. N.Y. Acad. Sci. 485, 73-79, 1986  
 A/Title: Recombinant genetic approaches to functional mapping of thrombin.  
 A/Reference number: I51952; M0ID:87182874; PMID:3471151  
 A/Accession: I51952  
 A/Status: translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-2, 'R', 5-100 <RES>  
 A/Cross-references: GB:M33031; NID:9190723; PIND:AAA60220.1; PID:9190724  
 C/Comment: Thrombin, which cleaves bonds after Arg and Lys, converts fibrinogen to fibrin and activates factors V, VIII, XIII, and, in complex with thrombomodulin, protein C.  
 C/Comment: Prothrombin is activated on the surface of a phospholipid membrane that binds the amino end of prothrombin and factors Va and Xa in calcium-dependent interactions. The activation peptide(s) can be removed either by factor Xa or thrombin; the cleavage into light and heavy chains is by factor Xa. It is not known whether one or two smaller activation peptides, with additional cleavage after 514-Arg, are released in natural blood clotting.  
 C/Comment: The cleavage after Arg-198, observed in vitro, does not occur in plasma.  
 C/Comment: The gamma-carboxyglutamic acid residues bind calcium ions, result from the carboxylation of glutamic acid residues by microsomal vitamin K-dependent carboxylase, and are necessary for calcium-dependent interaction with the negatively charged phospholipid membrane surface.  
 C/Comment: The prothrombin precursor is synthesized in the liver.  
 C/Genetics:

A:Gene: GDB:F2  
A:Cross-references: GDB:119894; OMIM:176930  
A:Map position: 11p11-11q12  
A:Introns: 27/1,86/3; 89/1; 106/1; 141/2; 187/1; 292/1; 335/1; 377/2; 433/2;  
491/2; 552/1; 575/3  
C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology  
C:Keywords: acute phase; blood coagulation; calcium binding; carboxylutamic  
acid; duplication; glycoprotein; hydrolase; kringle; liver; plasma; serine  
protease  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-43/Domain: propeptide #status predicted <PRO>  
F:48-87/Domain: Gla domain homology <GLA>  
F:44-622/Product: prothrombin #status experimental <NAT>  
F:44-327/Domain: activation peptide #status experimental <APT>  
F:108-186/Domain: kringle homology <KR1>  
F:213-291/Domain: kringle homology <KR2>  
F:328-363/Product: thrombin light chain #status experimental <LCH>  
F:364-622/Product: thrombin heavy chain #status experimental <HCH>  
F:364-613/Domain: trypsin homology <TRY>  
F:49/50,57,59,62,63,68,69,72,75/Modified site: gamma-carboxylutamic acid (Glu)  
#status experimental  
F:60-65,90-103,108-186,129-169,157-181,213-291,234-274,262-286/Disulfide bonds:  
#status predicted  
F:121,143/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:336-482,536-550,564-594/Disulfide bonds: #status predicted  
F:391-407/Disulfide bonds: #status experimental  
F:406,462/Active site: His, Asp #status predicted  
F:416/Binding site: carboxylate (Asn) (covalent) #status experimental  
F:568/Active site: Ser #status experimental

Query Match 100.0%; Score 131; DB 1; Length 622;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGYKPEEGKRGDAEGDSGSPFV 23  
|||||:|||||:|||||:|||||  
Db 551 AGYKPEEGKRGDAEGDSGSPFV 573

RESULT 2  
C42696  
thrombin (EC 3.4.21.5) B chain - rabbit (fragment)  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 26-May-1994 #sequence\_revision 26-May-1994 #text\_change 17-Mar-1999  
C:Accession: C42696  
R:Bantfield, D.K.; MacGillivray, R.T.A.  
Proc. Natl. Acad. Sci. U.S.A. 89; 2779-2783, 1992  
A:title: Partial characterization of vertebrate prothrombin cDNAs: amplification  
and sequence analysis of the B chain of thrombin from nine different species.  
A:Reference number: A42696; NUID:92212913; PMID:1557383  
A:Accession: C42696  
A:Status: preliminary; nucleic acid sequence not shown; not compared with  
conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-236 <BAN>  
A:Cross-references: GB:M61396  
C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology

C:Keywords: hydrolase; serine proteinase  
F:1-227/Domain: trypsin homology (fragment) <TRY>  
Query Match 96.9%; Score 127; DB 2; Length 236;  
Best Local Similarity 95.7%; Pred. No. 2.6e-10;  
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGYKPEEGKRGDAEGDSGSPFV 23  
|||||:|||||:|||||:|||||  
Db 165 AGYKPEEGKRGDAEGDSGSPFV 187

Search completed: February 11, 2004, 14:56:57  
Job time : 15.5806 secs

OM protein - protein search, using sw model

Run on: February 11, 2004, 14:36:52 ; Search time 9.64516 Seconds

(without alignments)  
112.141 Million cell updates/sec

Title: US-10-050-611-4

Perfect score: 131

Sequence: 1 AGYKPECKRGKDGACGDSGAPTV 23

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	131	100.0	622	1	THRB_HUMAN
2	124	94.7	625	1	THRB_BOVIN
3	102	77.9	617	1	THRB_RAT
4	102	77.9	618	1	THRB_MOUSE
5	73.5	56.1	290	1	MPN_HUMAN
6	71.5	54.6	417	1	HEPS_HUMAN
7	71.5	54.6	436	1	HEPS_MOUSE
8	71	54.2	161	1	PRTC_MACMU
9	71	54.2	461	1	PRTC_HUMAN
10	70.5	53.8	638	1	KAL_HUMAN
11	70	53.4	281	1	TRY2_DROER
12	69.5	53.1	275	1	TRY3_ANOGA
13	68.5	52.3	468	1	FA10_HUMAN
14	68.5	52.3	1019	1	LFC_CARRO
15	68.5	52.3	1019	1	LFC_TACTR
16	68	51.9	458	1	PRTC_RABIT
17	67.5	51.5	282	1	FA9_RAT

ALIGNMENTS

RESULT 1	THRB_HUMAN	STANDARD	PRT	622 AA.
AC	P00734;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-JAN-1990 (Rel. 13, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Prothrombin precursor (EC 3.4.21.5) (Coagulation factor II).			
GN	F2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88077877; PubMed=2825773;			
RA	Degen S.J.F., Davie E.W.;			
RT	"Nucleotide sequence of the gene for human prothrombin.";			
RL	Biochemistry 26:6165-6177(1987).			
RN	[2]			
RP	SEQUENCE FROM N.A., AND VARIANT MET-165.			
RA	Ridder M.O., Arnel T.Z., Cattington D.P., Chung M.-W., Lee K.L.,			
RA	Ozuna M., Poel C.L., Toth E.-J., Yi Q., Nickerson D.A.;			
RL	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.			

18	67.5	51.5	459	1	FA9_MOUSE	P16294	mus musculus
19	67.5	51.5	475	1	FA10_CHICK	P25155	gallus gall
20	67.5	51.5	638	1	KAL_MOUSE	P26262	mus musculus
21	67	51.1	256	1	KLKE_HUMAN	Q94245	homo sapien
22	67	51.1	264	1	VDP_BOMMO	Q07943	bombyx mori
23	66.5	50.8	455	1	TMS5_MOUSE	Q98040	mus musculus
24	66.5	50.8	457	1	TMS5_HUMAN	Q93363	homo sapien
25	65.5	50.0	267	1	TRY7_ANOGA	P35041	anopheles g
26	65.5	50.0	274	1	TRY1_ANOGA	P35035	anopheles g
27	65.5	50.0	275	1	TRY4_ANOGA	P35038	anopheles g
28	65.5	50.0	277	1	TRY2_ANOGA	P35036	anopheles g
29	65.5	50.0	638	1	KAL_RAT	P14272	rattus norv
30	65	49.6	157	1	PRTC_CARPA	Q28278	carls fami
31	65	49.6	157	1	PRTC_CAPHI	Q28315	capra hircu
32	65	49.6	157	1	PRTC_FELCA	Q28442	felis silve
33	65	49.6	157	1	PRTC_HORSE	Q28380	equus caball
34	65	49.6	459	1	PRTC_PIG	Q94182	sus scrofa
35	64.5	49.2	238	1	TRY5_AEDAE	P29787	aedes aegypt
36	64.5	49.2	422	1	DESI_HUMAN	Q9152	homo sapien
37	64.5	49.2	480	1	FA10_RABIT	O19045	oryctolagus
38	64	48.9	253	1	CPAD_HUMAN	P00746	homo sapien
39	64	48.9	259	1	CPAD_PIG	P31779	sus scrofa
40	64	48.9	456	1	PRTC_BOVIN	P00745	bos taurus
41	64	48.9	875	1	NETR_HUMAN	P6730	homo sapien
42	64	48.9	2616	1	NDL_DROER	P8159	drosophila
43	63.5	48.5	625	1	FA1_HUMAN	P03951	homo sapien
44	63	48.1	256	1	TRYE_DROER	P4627	drosophila
45	63	48.1	461	1	PRTC_MOUSE	P35587	mus musculus

RN [3]  
 RP SEQUENCE OF 8-622 FROM N.A.  
 RX MEDLINE=83231469; PubMed=6305407;  
 RA Degen S.J.F., McGillivray R.T.A., Davie E.W.;  
 RT "Characterization of the complementary deoxyribonucleic acid and gene  
 RT coding for human prothrombin.";  
 RL Biochemistry 22:2087-2097(1983).  
 RN [4]  
 RP SEQUENCE OF 44-314.  
 RX MEDLINE=77193964; PubMed=266717;  
 RA Walz D.A., Hewett-Emslett D., Seegers W.H.;  
 RT "Amino acid sequence of human prothrombin fragments 1 and 2.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 74:1969-1972(1977).  
 RN [5]  
 RP SEQUENCE OF 315-622.  
 RX MEDLINE=77207112; PubMed=873923;  
 RA Burkowski R.J., Eilon J., Downing M.R., Mann K.G.;  
 RT "Primary structure of human prothrombin 2 and alpha-thrombin.";  
 RL J. Biol. Chem. 252:4942-4957(1977).  
 RN [6]  
 RP PROCESSING.  
 RX MEDLINE=87008532; PubMed=3759956;  
 RA Rabiet M.J., Blaesill A., Furie B., Furie B.C.;  
 RT "Prothrombin fragment 1 X 2 X 3, a major product of prothrombin  
 RT activation in human plasma.";  
 RL J. Biol. Chem. 261:13210-13215(1986).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
 RX MEDLINE=90059942; PubMed=2583108;  
 RA Bode W., Mayr I., Baumann U., Huber R., Stone S.R., Hofsteenge J.;  
 RT "The refined 1.9 A crystal structure of human alpha-thrombin:  
 RT interaction with D-Phe-Pro-Arg chloromethylketone and significance of  
 RT the Tyr-Pro-Tyr insertion segment.";  
 RL EMBO J. 8:3467-3475(1989).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE=90327074; PubMed=2374926;  
 RA Rydel T.J., Ravichandran K.G., Tulinsky A., Bode W., Huber R.,  
 RA Rottsch C., Fenton J.W. II;  
 RT "The structure of a complex of recombinant hirudin and human alpha-  
 RT thrombin.";  
 RL Science 249:277-280(1990).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
 RX MEDLINE=94350942; PubMed=8071320;  
 RA Rydel T.J., Yin M., Pedraza-Naranjo K.P., Blankenship D.T., Cardin A.D.,  
 RA Correa P.E., Fenton J.W. II, Tulinsky A.;  
 RT "Crystallographic structure of human gamma-thrombin.";  
 RL J. Biol. Chem. 269:22000-22006(1994).  
 RN [10]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE=97357286; PubMed=9214615;  
 RA van de Locht A., Bode W., Huber R., le Bonniec B.F., Stone S.R.,  
 RA Esmen C.T., Stubbs M.T.;  
 RT "The thrombin E192Q-BPTI complex reveals gross structural  
 RT rearrangements: implications for the interaction with antithrombin  
 RT and thrombomodulin.";  
 RL EMBO J. 16:2977-2984(1997).  
 RN [11]  
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 328-601.  
 RX MEDLINE=99162521; PubMed=10051558;  
 RA Quinto E.R., Caccia S., Rose T., Fueterer K., Wakaman G., di Cera E.;  
 RT "Unexpected crucial role of residue 225 in serine proteases.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:1852-1857(1999).  
 RN [12]  
 RP VARIANT BARCELONA.  
 RX MEDLINE=87033739; PubMed=3771562;  
 RA Rabiet M.-J., Furie B.C., Furie B.;  
 RT "Molecular defect of prothrombin Barcelona. Substitution of cysteine  
 RT for arginine at residue 273.";  
 RL J. Biol. Chem. 261:15045-15048(1986).  
 RN [13]  
 RP VARIANT FRANKFURT.  
 RX MEDLINE=95313001; PubMed=7792730;  
 RA Degen S.J.F., McDowell S.A., Sparks L.M., Scharer I.;  
 RT "Prothrombin Frankfurt: a dysfunctional prothrombin characterized by  
 RT substitution of Glu-466 by Ala.";  
 RL Thromb. Haemost. 73:203-209(1995).  
 RN [14]  
 RP VARIANTS HIMI-1 AND HIMI-2.  
 RX MEDLINE=93043342; PubMed=1421398;  
 RA Morishita E., Saito M., Kumabashiri I., Asakura H., Matsuda T.,  
 RA Yamaguchi K.;  
 RT "Prothrombin Himi: a compound heterozygote for two dysfunctional  
 RT prothrombin molecules (Met-337->Phe and Arg-388->His).";  
 RL Blood 80:2275-2280(1992).  
 RN [15]  
 RP VARIANT PADUA-1.  
 RX MEDLINE=95169898; PubMed=7865694;  
 RA James H.L., Kim D.J., Zheng D.-Q., Girolami A.;  
 RT "Prothrombin Padua I: incomplete activation due to an amino acid  
 RT substitution at a factor Xa cleavage site.";  
 RL Blood Coagul. Fibrinolysis 5:841-844(1994).  
 RN [16]  
 RP VARIANT QUICK-1.  
 RX MEDLINE=89207504; PubMed=3242619;  
 RA Henriksen R.A., Mann K.G.;  
 RT "Identification of the primary structural defect in the dysfibrinogen  
 RT thrombin quick I: substitution of cysteine for arginine-382.";  
 RL Biochemistry 27:9160-9165(1988).  
 RN [17]  
 RP VARIANT QUICK-2.  
 RX MEDLINE=89247398; PubMed=2719946;  
 RA Henriksen R.A., Mann K.G.;  
 RT "Substitution of valine for glycine-558 in the congenital dysfibrinogen  
 RT thrombin quick II alters primary substrate specificity.";  
 RL Biochemistry 28:2078-2082(1989).  
 RN [18]  
 RP VARIANT SAIKITA.  
 RX MEDLINE=92378975; PubMed=1354985;  
 RA Miyata T., Aruga R., Uneyama H., Bezeaud A., Guillouin M.-C.,  
 RA Tanaga S.;  
 RT "Prothrombin Saikita: substitution of glutamic acid-466 by alanine  
 RT reduces the fibrinogen clotting activity and the esterase activity."

RL Biochemistry 31:7457-7462(1992).  
 RN [19]  
 RP VARIANT TOKUSHIMA.  
 RX MEDLINE=87185407; PubMed=3567158;  
 RA Miyata T., Morita T., Inomoto T., Kawachi S., Shirakami A.,  
 RA Iwanaga S.;  
 RT "Prothrombin Tokushima, a replacement of arginine-418 by tryptophan  
 RT that impairs the fibrinogen clotting activity of derived thrombin  
 RT Tokushima.";  
 RL Biochemistry 26:1117-1122(1987).  
 RN [20]  
 RP VARIANT TOKUSHIMA.  
 RX MEDLINE=87101511; PubMed=3801671;  
 RA Inomoto T., Shirakami A., Kawachi S., Shigeaki T., Saito S.,  
 RA Miyoshi K., Morita T., Iwanaga S.;  
 RT "Prothrombin Tokushima: characterization of dysfunctional thrombin  
 RT derived from a variant of human prothrombin.";  
 RL Blood 69:565-569(1987).  
 RN [21]  
 RP VARIANT TOKUSHIMA.  
 RX MEDLINE=92256895; PubMed=1349838;  
 RA Iwahana H., Yoshimoto K., Shigeaki T., Shirakami A., Saito S.,  
 RA Itakura M.;  
 RT "Detection of a single base substitution of the gene for prothrombin  
 RT Tokushima. The application of PCR-SSCP for the genetic and molecular  
 RT analysis of dysprothrombemia.";  
 RL Int. J. Hematol. 55:93-100(1992).  
 RN [22]  
 RP VARIANT TYPE-3.  
 RX MEDLINE=83204687; PubMed=6405779;  
 RA Board P.G., Shaw D.C.;  
 RT "Determination of the amino acid substitution in human prothrombin  
 RT type 3 (157 Glu leads to Lys) and the localization of a third  
 RT thrombin cleavage site.";  
 RL Br. J. Haematol. 54:245-254(1983).  
 RN [23]  
 RP VARIANTS MET-165 AND THR-386.  
 RX MEDLINE=99318093; PubMed=10391209;  
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.O.,  
 RA Lander E.S.;  
 RT "Characterization of single-nucleotide polymorphisms in coding regions  
 RT of human genes.";  
 RL Nat. Genet. 22:231-238(1999).  
 RN [24]  
 RP ERRATUM.  
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.O.,  
 RA Lander E.S.;  
 RL Nat. Genet. 23:373-373(1999).  
 CC -1- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS  
 CC FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII,  
 CC AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.  
 CC -1- CATALYTIC ACTIVITY: preferential cleavage: Arg-|-Gly; activates  
 CC fibrinogen to fibrin and releases fibrinopeptide A and B.

CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN THE LIVER; FOUND IN PLASMA.  
 CC -1- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,  
 CC RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOVAL  
 CC ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES  
 CC ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY  
 CC CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION  
 CC  
 Query Match 100.0%; Score 131; DB 1; Length 622;  
 Best Local Similarity 100.0%; Pred. No. 2,1e-10;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGKPDGKRGDACEGSGGPFV 23  
 DB 551 AGKPDGKRGDACEGSGGPFV 573  
 Search completed: February 11, 2004, 14:54:04  
 Job time : 9.64516 secs

OM protein - protein search, using sw model

Run on: February 11, 2004, 14:47:57 ; Search time 39.3226 Seconds  
(without alignments)  
150,936 Million cell updates/sec

Title: US-10-050-611-4

Perfect score: 131  
Sequence: 1 AGYKPEGRKEDACEGSGGPFV 23

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SPREMBL 23:\*
- 2: sp\_archaea:\*
- 3: sp\_bacteria:\*
- 4: sp\_fungi:\*
- 5: sp\_human:\*
- 6: sp\_invertebrate:\*
- 7: sp\_mammal:\*
- 8: sp\_mhc:\*
- 9: sp\_organelle:\*
- 10: sp\_phase:\*
- 11: sp\_plant:\*
- 12: sp\_podent:\*
- 13: sp\_virus:\*
- 14: sp\_vertebrate:\*
- 15: sp\_unclassified:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
---------------	-------	----------------	--------	----	-------------

1	127	96.9	235	6	Q28731	Q28731 oryctolagus
2	118	90.1	235	13	Q90387	Q90387 cynops pyrr
3	113	86.3	235	13	Q91004	Q91004 gekko gekko
4	113	86.3	607	13	Q91001	Q91001 gallus gall
5	113	86.3	608	13	Q9PTW7	Q9PTW7 struthio ca
6	109	83.2	239	13	Q91218	Q91218 oncorhynch
7	105	80.2	420	13	Q90504	Q90504 epistretus
8	98	74.8	172	13	Q9PFD1	Q9PFD1 oncorhynch
9	92	70.2	234	13	Q90244	Q90244 acipenser t
10	72.5	55.3	359	13	Q9PYX7	Q9PYX7 xenopus lae
11	72.5	55.3	974	13	Q90WD8	Q90WD8 bufo japoni
12	71.5	54.6	435	11	Q9C9Y7	Q9C9Y7 mus musculu
13	71.5	54.6	799	11	Q9DBT0	Q9DBT0 mus musculu
14	71.5	54.6	802	4	Q8IU80	Q8IU80 homo sapien
15	71.5	54.6	811	4	Q8IU80	Q8IU80 homo sapien
16	71	54.2	195	4	Q8U008	Q8U008 homo sapien
17	71	54.2	195	4	Q8U007	Q8U007 homo sapien
18	71	54.2	195	4	Q8U006	Q8U006 homo sapien
19	71	54.2	195	4	Q8IXB4	Q8IXB4 homo sapien
20	71	54.2	211	4	Q8U009	Q8U009 homo sapien
21	70.5	53.8	161	11	Q63109	Q63109 rattus norv
22	70.5	53.8	259	5	Q8XV61	Q8XV61 ctenecephal
23	70.5	53.8	267	5	Q9BK47	Q9BK47 lutula foli
24	70.5	53.8	481	11	Q54740	Q54740 mus musculu
25	70.5	53.8	481	11	Q99132	Q99132 mus musculu
26	70.5	53.8	481	11	Q88947	Q88947 mus musculu
27	70.5	53.8	482	11	Q63207	Q63207 rattus norv
28	70	53.4	378	5	Q8SV50	Q8SV50 drosophila
29	69.5	53.1	200	11	Q92406	Q92406 mus musculu
30	69.5	53.1	1524	13	Q91674	Q91674 xenopus lae
31	68.5	52.3	161	6	Q28511	Q28511 macaca mula
32	68.5	52.3	236	5	Q9TVH3	Q9TVH3 schistosoma
33	68.5	52.3	488	5	Q9TVH4	Q9TVH4 schistosoma
34	68.5	52.3	766	4	Q8NBY4	Q8NBY4 homo sapien
35	68.5	52.3	1019	5	Q8TP91	Q8TP91 tachypleus
36	68.5	52.3	1083	5	Q26423	Q26423 catclinoscor
37	68	51.9	686	13	Q9DGC2	Q9DGC2 cyprinus ca
38	67.5	51.5	156	5	Q16007	Q16007 schistosoma
39	67.5	51.5	161	11	Q60546	Q60546 mesocricetu
40	67.5	51.5	264	5	Q02569	Q02569 culx quing
41	67.5	51.5	328	11	Q8BJR6	Q8BJR6 mus musculu
42	67.5	51.5	370	5	Q9V444	Q9V444 drosophila
43	67.5	51.5	387	5	Q8XV57	Q8XV57 ctenecephal
44	67.5	51.5	474	13	Q8UJ08	Q8UJ08 brachydanto
45	67.5	51.5	638	11	Q8R0P5	Q8R0P5 mus musculu

Search completed: February 11, 2004, 14:56:05  
Job time : 39.3226 secs